Highly Efficient and Reliable Quantum Cascade Lasers

By

Benjamin Knipfer

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Electrical and Computer Engineering)

at the
UNIVERSITY OF WISCONSIN-MADISON
2021

Date of final oral examination: 12/20/2021

The dissertation is approved by the following members of the Final Oral Committee:

Dan Botez, Professor, Electrical and Computer Engineering
Luke J. Mawst, Professor, Electrical and Computer Engineering
Susan Babcock, Professor, Materials Science and Engineering
Irena Knezevic, Professor, Electrical and Computer Engineering

Abstract

Mid-infrared (mid-IR) quantum cascade lasers (QCLs) have been commercially available for low power applications, however, while the desire for higher power devices is present, the efficiency and reliability are severe limitations. This work takes a multi-faceted approach to improving the reliability and efficiency of QCLs including: the identification and mitigation of failure mechanisms under high power continuous wave (CW) and quasi-continuous wave (QCW) operation, optical and thermal modeling of devices to further reduce active region heating, verification of these models using charge-coupled device (CCD) based thermoreflectance, and the introduction of interface roughness (IFR) engineered devices to reduce IFR scattering and leakage. Atom probe tomography (APT) is also employed to investigate the amount of aluminum and gallium incorporation in thin InAlAs barriers and InGaAs wells. It was found that thin layers with thicknesses less than 2 nm require an intentional aluminum or gallium overshoot in the gas phase during growth to grow the targeted compositions. This was verified when the overshoot in thin barriers resulted in the convergence of modeled and experimental emitting wavelengths. APT was also used to interrogate a few key interfaces within a 40 stage strain-compensated QCL emitting near 4.6 μm. This interrogation yielded both in-plane and axial IFR parameters for barriers of high and low aluminum incorporation, and in turn high and low strain, respectively. It was found that the barrier with the highest aluminum target had a nearly 50% larger root mean square (RMS) roughness when compared to the shorter barriers. As the IFR scattering is proportional to the square of both the RMS roughness and in-plane correlation length, this finding has a significant impact on the IFR scattering and leakage. The variable IFR parameters, axial correlation length, graded interfaces, graded lattice constants, graded conduction band edge, and quaternary alloy disorder (AD) scattering have been incorporated into a scattering model. Results from this model suggest lower global lifetimes and significantly reduced

transition efficiencies which results in lower IFR leakage, however, if electronic temperatures from software using non-equilibrium Green's function (NGEF) is incorporated, leakage currents remain high.

Acknowledgements

First and foremost, I would like to thank my advisors, Prof. Mawst and Prof. Botez; without their guidance none of this would have been possible. I would also like to thank my preliminary exam and defense committee members, Prof. Irena Knezevic and Prof. Susan Babcock, for their time and direction.

I would like to acknowledge previous and current group members: Dr. Chris Sigler, Dr. Ayushi Rajeev, Dr. Honghyuk Kim, Dr. Colin Boyle, Dr. Kevin Oresick, Dr. Jeremy Kirch, Jae Ha Ryu, Shining Xu, Suraj Suri, Huilong Gao, Shuqi Zhang, and Morgan Turville-Heitz. Thank you all for all your help - your companionship and support were essential in this process.

I would like to thank Intraband employees, previous and current, Tom Earles, Don Lindberg III, Steven Jacobs, Dave Hoerr, and Rob Marsland for their insights and new perspectives.

Furthermore, I would like to extend thanks to the staff at the Nanoscale Fabrication Center (NFC) for training me on various tools and their respective expertise in adjusting processes and conditions to achieve my desired outcomes. I would also like to thank the staff at the Nanoscale Imaging and Analysis Center (NIAC) for their training as well, specifically with the in-depth knowledge and use of the focused ion beam tools. I would like to additionally thank Dr. Dieter Isheim at Northwestern University Center for Atom Probe Tomography (NUCAPT) whose collaboration and APT expertise was invaluable.

Most notably I would like to thank my friends and family, especially my partner, Emeline Beck, for their unending support and encouragement.

Table of Contents

JOURNAL PUBLICATIONS, CONFERENCE PRECEDINGS, INVITED SEMINARS, AND CONFERENCE PRESENTATIONS

B. Knipfer, S. Xu, J. D. Kirch, D. Botez, L. J. Mawst, "Analysis of Interface roughness in Strained InGaAs/AlInAs Quantum Cascade Laser Structures (λ ~ 4.6 um) by Atom Probe Tomography", J. Crys. Growth (2022) [under review]

S. Suri, **B. Knipfer**, J. D. Kirch, L. J. Mawst, T. Grange, D. Botez, "Rigorous modeling of mid-IR QCLs with strong photon-induced carrier transport", SPIE, Photonics West (2022)

B. Knipfer, A. Rajeev, D. Isheim, J. D. Kirch, S. E. Babcock, T. Kuech, T. Earles, D. Botez, L. J. Mawst. "Layer Thickness Dependent Interfacial Grading Analysis for Strained III-V Superlattices by Atom Probe Tomography" J. Crys. Growth. 535, 125550 (2020)

Y. Sin, Z. Lingley, M. Brodie, **B. Knipfer**, C. Sigler, C. Boyle, J. D. Kirch, K. Oresick, H. Kim, D. Botez, L. J. Mawst, D. Lindberg III, and T. Earles. "Degradation Mechanisms in MOVPE-Grown High-Power Buried-Heterostructure Quantum Cascade Lasers" – CLEO 2020

B. Knipfer, A. Rajeev, D. Isheim, J. D. Kirch, S. E. Babcock, T. Kuech, T. Earles, D. Botez, L. J. Mawst. "Layer Thickness Dependent Interfacial Grading Analysis for Strained III-V Superlattices by Atom Probe Tomography" at International Conference on Crystal Growth and Epitaxy (ICCGE-19) 2019

N. Becher, M. Farzaneh, **B. Knipfer**, C. Sigler, J. Kirch, C. Boyle, D. Botez, L.J. Mawst, D. F. Lindberg III, and T. Earles. "Thermal Imaging of Buried Heterostructure Quantum Cascade Lasers (QCLs) and QCL Arrays using CCD-based Thermoreflectance Microscopy". Journal of Appl. Phys. Vol. 125,3 (2019)

D. Botez, J. D. Kirch, C. Boyle, K. M. Oresick, C. Sigler, H. Kim, **B. Knipfer**, J. H. Ryu, D. F. Lindberg, T. L. Earles, L. J. Mawst, and Y. V. Flores, "High-efficiency, high-power mid-infrared quantum cascade lasers [Invited]," Opt. Mater. Express8, 1378-1398 (2018).

B. Knipfer, C. Sigler, C. Boyle, J. D. Kirch, K. Oresick, H. Kim, D. Botez, L. J. Mawst, N. Becher, M. Farzaneh, and D. F. Lindberg, "Failure Analysis of High-Power (One-Watt) Room-Temperature Continuous Wave MOCVD Quantum Cascade Lasers," 2018 IEEE International Semiconductor Laser Conference (ISLC), IEEE, 2018

Y. Sin, Z. Lingley, M. Brodie, **B. Knipfer**, C. Sigler, C. Boyle, J. D. Kirch, K. Oresick, H. Kim, D. Botez. L. J. Mawst, D. Lindberg, and T. Earles "Catastrophic Degradation in High-Power Buried Heterostructure Quantum Cascade Lasers" Presented at CLEO – 2019 by the AFRL

C. Boyle, K. Oresick, J. D. Kirch, C. Sigler, **B. Knipfer**, L. J. Mawst, D. F. Lindberg, T. Earles, Y. V. Flores, and D. Botez, "95 % Laser-Transition Efficiency from Step-Taper Active-Region Mid-IR QCLs," at International Quantum Cascade Lasers School and Workshop (IQCLSW), 2018

List of Figures

aluminum group III ratio, **d)** 20% aluminum group III ratio, **e)** 21% aluminum group III ratio, **f)** 22% aluminum group III ratio, **g)** 23% aluminum group III ratio, and **h)** 24% aluminum group III ratio.

Figure 6.30 Components of leakage for 5g_mod when using an axial correlation length of 0.1 nm, an interfacial mixing width of 0.5 nm, Δ and Λ_{II} are being varied from 0.13 nm to 0.2 nm and 6.2 nm to 7.1 nm, respectively, from the APT extracted values, and quaternary AD scattering.

List of Tables

CHAPTER ONE

INTRODUCTION

1.1 Introduction

Quantum cascade lasers (QCLs) are unipolar semiconductor lasers that employ alternating layers of InAlAs and InGaAs of varying thicknesses and compositions to achieve wavelength agility, as opposed to having the emitting wavelength being directly tied to the band gap of any of the constituent materials, and, through active region design optimization, QCLs have significant potential for low carrier leakage, device temperature insensitivity, high efficiency, and high output power. These devices utilize intersubband transitions from quantized states within the conduction band that are engineered through quantum confinement. High-power QCLs emitting in the mid-infrared (mid-IR) have secured much interest for their applications in remote sensing, IR countermeasures, and long-range communications. The mid-IR is an advantageous wavelength regime as there are two main windows where there is low atmospheric absorption: 3-5 μm, and 8-12 μm. Emission in both these windows can be achieved through QCL active region design, but for the purposes here the focus will be on 3-5 μm emitting wavelength devices. The other main mid-IR sources have typically been lead salt lasers, whose highest temperature to date for CW operation is only 223 K [1] and interband cascade lasers (ICLs). ICLs can reach similar wavelengths as QCLs with lower current densities, however, they also have much lower power densities. State-of-the-art ICLs exhibit peak room temperature CW power <500 mW as a result of high Auger losses and internal losses [2,3]. As QCLs use intersubband transitions they do not experience the same Auger losses that their band-to-band counterparts exhibit and thus their threshold current density dependence on temperature is much lower [4].

The typical QCL has three regions: the injector, the active, and the extractor regions. The injector acts as a reservoir for electrons that are injected into the active region where the optical transition occurs, and an efficient extractor region, where electrons are extracted, is necessary to create the essential carrier population inversion. Electrically pumped interband devices naturally have a population inversion across the band gap, but QCLs need to be designed such that there is a strong population inversion by creating relatively long upper-level lifetimes and minimizing the lower-level lifetimes through efficient extraction of carriers.

The optical transition has a relatively low probably of emitting a photon when compared to their interband counterparts; to make functional, efficient devices a period must be repeated 20-60 times. This presents the opportunity for one electron to emit up to the number of repetitions of the period, also commonly referred to as a stage, as shown in Figure 1.1.

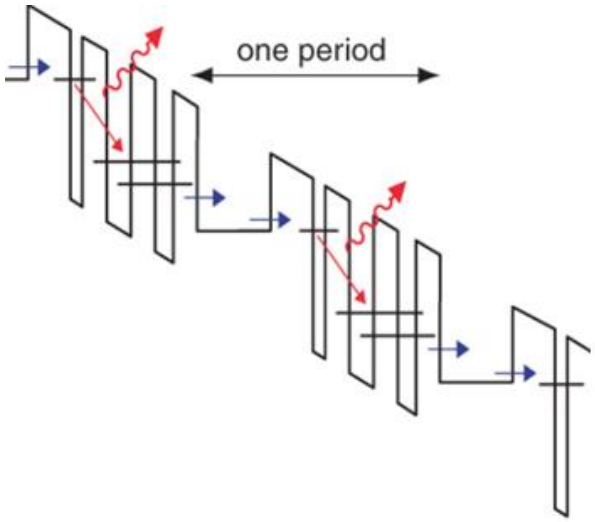


Fig. 1.1 A simplistic image showing the conduction band of part of a QCL active region. The solid horizontal lines show the electronic states, the blue arrows show the extraction and injection of electrons, the red arrows signify the lasing transition, and the red arrow pointing up and right signifies the emission of a photon with a designed characteristic wavelength.

Each of these stages contain on the order of ~20 layers which results in roughly 400-1200 layers within the superlattice (SL) for a device with 20-60 stages. The first experimentally demonstrated QCL was in 1994 at Bell Labs [5]. These first devices were grown via molecular beam epitaxy (MBE) with an emitting wavelength of 4.2 μm and a peak power of ~8mW. In 2002, the first continuous wave (CW) operation at room temperature was shown [6]. Through active region development [7], growth optimization [8,9], and thermal management strategies, wavelength is highly controllable and peak CW powers can reach in excess of 5W CW at room temperature with wall-plug efficiencies (WPE) greater than 25% [11]. Thermal management is such a key parameter for these devices that the waveguide design has included a spacer, made of InP, within the active region and the growth of Fe-doped InP around the ridge for lateral heat removal as well as current confinement. As it is such a critical parameter some have begun to use it as a method of tuning wavelength [10].

1.2 Thesis Overview

This thesis focuses on the advancement towards highly efficient and reliable QCLs emitting in the mid-IR grown via metal organic chemical vapor deposition (MOCVD). There are many variables at play including: self-heating based off device structure and geometry and the role it plays on device performance, optical losses, and carrier leakage. As this is such a complex problem a multi-faceted approach is utilized. The device reliability component of this work is addressed through the mitigation of observed failure mechanisms. Thermal and optical modeling is employed as the efficiency of these devices is dependent on the optical losses and device heating. The other aspect of design that has the largest impact on device efficiency is the active region design. To address this area of potential efficiency

increase, and to further optimize active region designs, the scattering rates are characterized and are included in a scattering model.

The focus of Chapter 2 is on the various scattering mechanisms within a QCL active region. The discussed scattering mechanisms will include longitudinal optical phonon scattering, alloy disorder scattering, and interface roughness scattering and how these different mechanisms relate to the active region design.

In Chapter 3, the investigation of QCL failure mechanisms is performed. The failure mechanisms for other edge emitters such as diodes are well understood and well-documented. Diodes typically fail at a critical input power density (or optical power density) and while QCLs are also sensitive to the amount of input power, the failure mechanisms are different from those of diode lasers. After the identification and characterization of these failure modes, various solutions are implemented to either solve or mitigate these issues. This work motivates further improvement of heat management and device performance, which is the subject of subsequent chapters.

In Chapter 4, the thermal and optical properties of the waveguide, and various waveguide configurations, are investigated using COMSOL Multiphysics. With thermoreflectance measurements the Multiphysics model is verified, and the thermal resistance of a device is quantified [12]. As thermal dissipation is a key element of waveguide design, various waveguide geometries are analyzed for optimal heat dissipation. The other aspect of thermal, and in turn optical, properties is the temperature dependence of the active region itself – this chapter will also investigate the T_0T_1 values for a few different designs and an empirical model to estimate these values for a potential QCL active region. The T_0T_1 values coupled with the experimental pulsed light-current-voltage (LIV) curves can be used to estimate the CW performance of a device with a given number of stages, dimensions, mirror reflectivities, and thermal resistance.

Chapter 5 will focus on carrier leakage through different scattering mechanisms, the resulting effect on device performance, and how active regions have been designed to combat these issues. A novel method of designing active regions is introduced: IFR-Engineered designs. A variation of a previously grown design using IFR engineering is examined and shows how further improvements in leakage suppression and internal efficiencies can be realized.

Chapter 6 will focus mainly on work performed with atom probe tomography (APT). Previous APT work related to QCL's is analyzed. APT is used to look at thin (<2nm) InAlAs barriers and aluminum incorporation relative to the target growth thickness. These findings are incorporated into a model and the relation to experiment is explored. The focus will then shift back to IFR scattering and how previous studies have used APT to extract IFR parameters. Relevant IFR parameters for a mid-IR QCL grown via MOCVD are extracted and are found to vary with different barrier heights, or aluminum concentration in InAlAs barriers, and in turn strain. Variable IFR parameters are then implemented in a scattering model. The results from this scattering model, including threshold current density, slope, and internal efficiency, are correlated with experimental values.

Chapter 7 will summarize the key findings and outline future work that can further expand these studies and this work towards more reliable and efficient QCLs emitting in the mid-infrared (mid-IR).

References

[1] Z. Feit, M. McDonald, R. J. Woods, V. Archambault, and P. Mark, "Low threshold PbEuSeTe/PbTe separate confinement buried heterostructure diode lasers", Appl. Phys. Lett. **68**, 738 (1996)

[2] W. Bewley, C. L. Canedy, C. S. Kim, M. Kim, C. D. Merritt, J. Abell, I. Vurgaftman, and J. Meyer, "Continuous-wave interband cascade lasers operating above room temperature at λ ~ 4.7-5.6 μm", Opt. Exp. Vol. 20, No. 3, pp. 3235-3240 (2012)

[3] J. R. Meyer, W. W. Bewley, C. L. Canedy, C. S. Kim, M. Kim, C. D. Merritt, and I. Vurgaftman, "The Interband Cascade Laser", Photonics 2020, 7(3), 75 (2020)

[4] E. Kioupakis, Q. Yan, D. Steiauf, and C. G. Van de Walle, "Temperature and carrier-density dependence of Auger and radiative recombination in nitride optoelectronic devices", New J. Phys. **15** 125006 (2013)

[5] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum cascade laser," Science, Vol. 264, No. 5158, pg. 553, (1994)

[6] D. Hofstetter, M. Beck, T. Aellen, and J. Faist, "Room temperature continuous wave operation of quantum cascade lasers", 5th International Conference on Mid-Infrared Optoelectronics Materials and Devices, 8–11 September, Annapolis, Maryland (2002)

[7] D. Botez, J. D. Kirch, C. Boyle, K. M. Oresick, C. Sigler, H. Kim, B. B. Knipfer, J. H. Ryu, D. Lindberg, T. Earles, L. J. Mawst, and Y. V. Flores. "High-efficiency, high-power mid-infrared quantum cascade lasers [Invited], Opt. Mater. Exp. Vol. 8, No. 5, 1378, (2018)

[8] P. Gutowski, I. Sankowski, T. Slupinski, D. Pierscinska, K. Pierscinski, A. Kuzmicz, K. Golaszewska-Malec, and M. Bugajski, "Optimization of MBE growth conditions of $In_{0.52}Al_{0.48}As$ waveguide layers for InGaAs/InAlAs/InP Quantum Cascade Lasers", Materials (Basel) 12(10): 1621 (2019)

[9] Y. Huang, J. H. Ryou, R. Dupuis, C. Pfluegl, F. Capasso, K. Sun, A. M. Fischer, and F. A. Ponce,"Optimization of growth conditions for InGaAs/InAlAs/InP quantum cascade lasers by metalorganic chemical vapor deposition", J. Crystl. Growth, 316(1):75-80 (2011)

[10] A. Muller, S. Blaser, J. Butet, T. Gresch, R. Maulini, and N. Villa, "Thermal management of packaged QCL", Swiss Photonics Workshop, (2019)

[11] Y. Bai, N. Bandyopadhyay, S. Tsao, S. Slivken, and M. Razeghi, "Room temperature quantum cascade lasers with 27% wall plug efficiency", Appl. Phys. Lett. 98, 181102 (2011)

[12] N. Becher, M. Farzaneh, B. Knipfer, C. Sigler, J. D. Kirch, C. Boyle, D. Botez, L. J. Mawst, D. F. Lindberg III, and T. Earles, "Thermal imaging of buried heterostructure quantum cascade lasers (QCLs) and QCL arrays using CCD-based thermoreflectance microscopy", J. of Appl. Phys. **125**, 033102 (2019)

CHAPTER TWO

BACKGROUND ON MID-INFRARED QCLs

2.1 Introduction

QCLs, which take advantage of intersubband transitions, have a distinct advantage over interband-transition semiconductor lasers in the mid-IR since there is little to no Auger recombination involved in intersubband transitions. However, the scientific community was initially surprised these devices even lased considering the extremely low upper-level lifetimes which are in the range of picoseconds. The low lifetimes are a result of various scattering mechanisms inherent to intraband transitions between energy states in multi-quantum-well structures. However, until recently, it has been commonly accepted that longitudinal-optical (LO) phonon scattering plays the dominant role in the carrier lifetimes for QCLs. Recent studies have shown that the gap between the theoretical and experimental internal-efficiency values can be bridged only through the inclusion of interface roughness (IFR) scattering and alloy disorder (AD) scattering in addition to LO-phonon scattering.

2.2 Longitudinal Optical Phonon Scattering

Phonons are described as lattice vibrations which are commonly assumed to follow an Arrhenius dependency with temperature. At non-zero temperatures lattice vibrations affect carrier lifetimes. It has been believed that QCL carrier lifetimes are dominated by LO-phonon scattering and as such many studies today consider only this scattering mechanism. LO scattering is an inelastic scattering event that involves the transfer of momentum either from or to an LO phonon, which is described as the interaction of the

electron with the deformed lattice and the resulting piezoelectric potential [1]. Quantitatively the momentum can be described as:

$$Q = \sqrt{k_i^2 + k_f^2 - 2k_i k_f \cos\,\theta} \tag{2.0}$$

where k_i is the initial wavevector, k_f is the final wavevector, and θ is the angle between the two wavevectors. This first approach, shown in equation 2.0 above, was outlined by Price [2]. This formalism uses the assumption that there is no LO phonon dispersion and the LO phonon scattering is dominated by the Frohlich Hamiltonian. With these two assumptions, the scattering rate can be written as:

$$W_{LO,scattering} = \frac{m^* e^2 \omega_{LO}}{2\hbar^2 \epsilon_P} \sum_f \int_0^{2\pi} d\theta \frac{I_{ij}(Q)}{Q} \tag{2.1}$$

where the inputs to this equation are defined below.

$$I_{ij}(Q) = \int dz \int dz' \varphi_i(z)\varphi_j(z) e^{-Q|z-z'|} \varphi_i(z')\varphi_j(z') \tag{2.2}$$

$$\epsilon_P^{-1} = \epsilon_\infty^{-1} - \epsilon_s^{-1} \tag{2.3}$$

$I_{ij}(Q)$ is the wavefunction envelope, the ϵ_∞^{-1} term is the mid-IR dielectric constant, ϵ_s^{-1} is the static dielectric constant, and ω_{LO} is the angular frequency of an LO phonon.

The above equations are considering carriers emitting an LO phonon and transitioning to a lower energy state, however, the inverse can be true as well. A carrier can receive the momentum transfer and be excited to a higher energy level. This carrier can then either be lost to the continuum, relax to the lower laser level, or a lower energy state below the lower laser level. Carriers that bypass the lasing transition act as a shunt-type leakage. The figure below shows these transitions schematically [3].

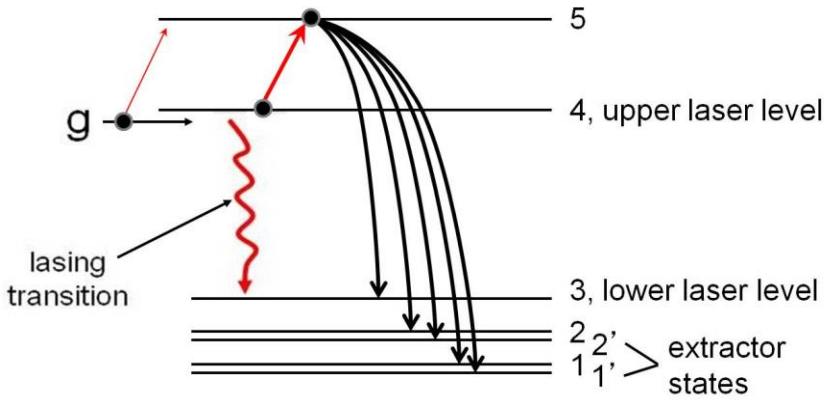


Fig. 2.1 Schematic drawing of a carrier excited to state 5 from either the ground state (labeled as 'g') or the upper laser (ul) level, 4. This carrier can then bypass the lasing transition and transition to the lower laser level or the states below, shown as the solid black arrows [3].

2.3 Alloy Disorder Scattering

Alloy disorder (AD) scattering, which is an elastic scattering event, is another mechanism that must be included for accurate modeling of devices. For the purposes of mid-IR QCLs, the material system is a ternary system that is alternating layers of InAlAs and InGaAs with specific thicknesses and compositions. From the perspective of AD scattering, these ternary systems are viewed as a binary with a third species as an interstitial or impurity described as $A_xB_{1-x}C$. The scattering rate between two states *i* and *j* , within a layer of given alloy, can be calculated using the following equation [4][5]:

$$\frac{1}{\tau_{ij}^{AD}} = \frac{1}{8}\frac{m^* a^3 (\delta E_c)^2 f_A f_B}{\hbar^3} \int_{alloy} dz\ \varphi_i^2(z)\varphi_j^2(z) \qquad (2.4)$$

where δE_c is the difference between the conduction-band (CB) minima of the well/barrier binary-alloy components, a is the lattice spacing, m* is the effective mass of the well/barrier, and $f_A$and f_B are the alloy fractions for their respective elements. It is evident the AD lifetime will reach a minimum when the

alloy fractions are both 0.5. Unfortunately, the lattice matched to InP compositions for InGaAs and InAlAs are $In_{0.53}Ga_{0.47}As$ and $In_{0.52}Al_{0.48}As$ which are both near the maximum AD scattering rates.

However, recent APT work [6] has shown that the interfaces are better defined as quaternaries with an extracted amount of intermixing. There is no published work involving the alloy disorder scattering for an InAlGaAs quaternary, however, the same method from Li [7], can be applied. Li's method has been created for how the alloy disorder scattering rate Impacts the mobility of a two-dimensional electron gas in AlInGaN, which has 3 group III elements and 1 group V element to match this material system, and the formalism does not consider the specific lattice structure. As such, the application of the same approach to this material system is justified. The scattering rate for a $A_xB_yC_{1-x-y}D$ is shown in equation 2.5 below.

$$\begin{aligned}\frac{1}{\tau^{AD}} = \frac{m^* a^3}{\hbar^3}\{&x(1-x-y)(1-y)(\Delta V_{Al-Ga})^2 \\ &+ y(1-x-y)(1-x)(\Delta V_{In-Ga})^2 \\ &+ xy(x+y)(\Delta V_{Al-In})^2 \\ &+ 2xy(1-x-y)(\Delta V_{Al-Ga}\Delta V_{In-Ga} \\ &+ \Delta V_{Al-Ga}\Delta V_{Al-In} - \Delta V_{In-Ga}\Delta V_{Al-In})\}\frac{\kappa_b P_b^2}{2}\end{aligned} \tag{2.5}$$

Where m^* is the effective mass, a^3 is the volume of the unit cell, assuming a cubic structure, k_b is a variational parameter which is proportional to $\Delta E_c^{1/2}$, $P_b = (Nz_0)^2/k_b$

The proportions of the binary compounds are then taken from the above equation and applied to equation 2.4, resulting in the following equation for the electron scattering rate in $Al_xIn_yGa_{1-x-y}As$:

$$\frac{1}{\tau_{ij}^{AD}} = \frac{1}{8}\frac{m^* a(z)^3}{\hbar^3}[x(1-x-y)(1-y)(\Delta V_{Al-Ga})^2 + y(1-x-y)\,(1-$$

$$x)(\Delta V_{In-Ga})^2 + xy(x+y)(\Delta V_{Al-In})^2 + 2xy\,(1-x-y)(\Delta V_{Al-Ga}\Delta V_{In-Ga} + \quad (2.6)$$

$$\Delta V_{Al-Ga}\Delta V_{Al-In} - \Delta V_{In-Ga}\Delta V_{Al-In})] \int_{alloy} dz\, \varphi_i^2(z)\varphi_j^2(z)$$

where,

$$\Delta V_{Al-Ga} = 0.8\ eV$$

$$\Delta V_{Al-In} = -0.6\ eV$$

$$\Delta V_{In-Ga} = 1.4\ eV$$

Incorporation of this new quaternary AD scattering for a 4.6 μm-emitting STA QCL, to be discussed in Chapter 6, results in a slightly increased upper-level global AD lifetime from 1.93 ps to 2.02 ps, but significantly increased the lower-level global AD lifetime from 4.41 ps to 6.07 ps.

2.4 Interface Roughness Scattering

Regardless of the growth method, whether it is molecular beam epitaxy (MBE), gas source MBE (GSMBE), or MOCVD, atomically sharp interfaces between wells and barriers cannot be achieved which leads to interface roughness (IFR) scattering and the main source of linewidth broadening as shown by Unuma [8,9]. The linewidth broadening can be calculated using the following equation:

$$2\Upsilon_{ij} = \frac{\pi m^*}{\hbar^2}\Delta^2\,\Lambda^2\delta U^2(z_k)\Sigma_k\left[\varphi_i^2(z_k) - \varphi_j^2(z_k)\right]^2 \quad (2.7)$$

IFR has been characterized using two parameters, Δ and Λ. These two parameters are schematically shown in Figure 2.2 for an arbitrary interface.

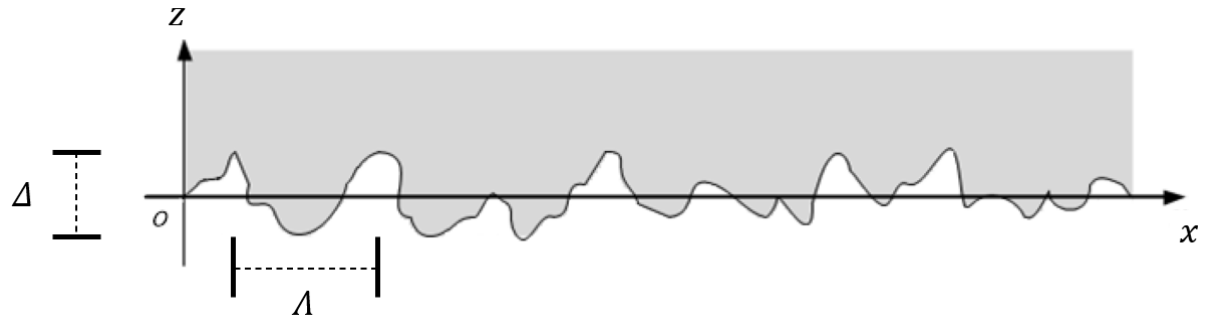


Figure 2.2 Schematic drawing of an arbitrary rough interface to show qualitatively the meaning of Λ and Δ IFR parameters.

The RMS roughness parallel to the growth direction is described as Δ, and Λ is the in-plane correlation length. The interface is assumed to have a Gaussian distribution and can be fit using the following auto-correlation function:

$$\langle \Delta(r)\Delta(r')\rangle = \Delta^2 \exp\left(-\frac{|r-r'|^2}{\Lambda^2}\right) \tag{2.8}$$

r is a vector in the (x,y) plane and r' is a vector with an arbitrary position while remaining in the (x,y) plane. As mentioned previously, although no growth method can achieve a perfect interface, some techniques yield lower roughness values. MBE is found to have values of about 0.1 nm and 9 nm for Δ and Λ respectively. GSMBE has slightly rougher interfaces with values of 0.125 nm and 9.3 nm. Of the 3 mentioned growth techniques, MOCVD, while industrially available and allows for much higher throughput, does have the roughest surface ranging from 0.12 nm – 0.13 nm for the Δ value and from 12.4 nm – 15 nm for the Λ value, for 4.6-5.0 μm-emitting QCLs. The IFR scattering rate between states n and m, where n is a state above m, at the i^{th} interface is calculated using the equation below [10]:

$$\left.\frac{1}{\tau_{nm}^{IFR}}\right|_i = \frac{\pi m^*}{\hbar^3}\Delta^2\Lambda^2\delta U_i^2(z_i)\varphi_n{}^2(z_i)\varphi_m{}^2(z_i)e^{-\frac{\Lambda^2 m^* E_{nm}}{2\hbar^2}} \tag{2.9}$$

This scattering calculation is performed between two wavefunctions for every interface within one stage of a QCL active region and then summed for a single scattering rate between two states as shown in equation 2.10.

$$\frac{1}{\tau_{nm}^{IFR}} = \sum_i \frac{1}{\tau^{IFR}}\Big|_i \tag{2.10}$$

This is done between every state within one stage of an active region to find scattering rates between every state up to two states above the upper laser (*ul*) level (i.e., up to state *ul*+2). These equations, in conjunction with LO and AD scattering, can be used to calculate global lifetimes for all states within an active region and the leakage of carriers can be analyzed. The figure below shows how carrier up-scattering can take carriers from the *ul* level, the g, or g_1 state, to the *ul*+1 state, and from there can either return to a level that can still participate in lasing or bypass the lasing transition [5].

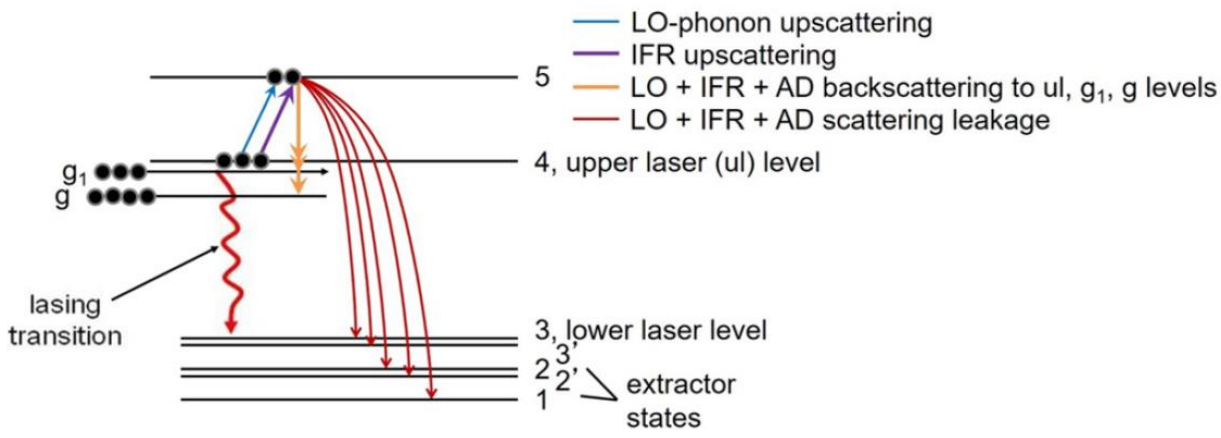


Fig. 2.3 Simplified schematic outlining the various carrier paths. Carriers can up-scatter from states 4, g_0, and g_1, to state 5. Carriers in state 5 can then relax back down to states 4, g and g_1 and still participate in the lasing transition or relax to a state below the lower laser (*ll)* level and act as a shunt type of leakage.

Flores was able to quantify the upward IFR leakage by using the following equations from state n to state m [11]

$$J_{nm}^{IFR} = \frac{en_n}{\tau_{mn}^{IFR}} I_{mn}(x_o) e^{-x_o} \tag{2.11}$$

where,

$$I_{mn}(x_o) = \int_0^{\infty} \exp\left(a_{mn} \left(\left[\frac{x}{x_o}\right]^2 + \frac{x}{x_o} \right)^{0.5} - x\left(1 + \frac{a_{mn}}{x_o}\right) \right) dx \tag{2.12}$$

In Equation 2.11 and 2.12, n_n is the carrier sheet density in state n, $a_{mn} = \mathrm{m}^* \Lambda^2 E_{mn}/\hbar^2$, and x_o = $E_{mn}/k_B T_{e,i}$. However, Flores et al does not consider carriers backscattered to ul, g_0, or g_1 levels that can still participate in the lasing transition. Boyle et al has since added this term for the IFR leakage calculation and has been shown to bridge the gap between the experimental and theoretical internal-efficiency values [5]. The updated equation is shown below in Equation 2.13 where $\tau_{n,tot}^{LO,IFR,AD}$ is the global lifetime of state n to every state below it, and $\tau_{n,leak}^{LO,IFR,AD}$ is the lifetime from state n to all low-energy active region and extractor states which takes into account the carriers backscattered to *ul*, g_0, and g_1, that can still participate in the lasing transition.

$$J_{nm}^{IFR} = \frac{en_n}{\tau_{mn}^{IFR}} \frac{\tau_{m,tot}^{LO,IFR,AD}}{\tau_{m,leak}^{LO,IFR,AD}} I_{mn}(x_o) e^{-x_o} \tag{2.13}$$

The inclusion of IFR and LO scattering also leads to an alteration of the leakage current calculation for LO-triggered leakage; this equation can be found below in Equation 2.14.

$$J_{leak,nm}^{LO} = \frac{en_n}{\tau_{mn}^{IFR}} \frac{\tau_{m,tot}^{LO,IFR,AD}}{\tau_{m,leak}^{LO,IFR,AD}} e^{-x_o} \tag{2.14}$$

These IFR calculations under the assumption the roughness within each layer is fully correlated. However, recent studies have experimentally determined this assumption is invalid and prompted the creation of a more general theoretical framework [12,13]. This framework brings two additional parameters into account for these graded interfaces, the interfacial mixing width L, and the axial correlation length, $\Lambda_\perp$. The extraction of these parameters is discussed in Chapter 5. The new formalism for the IFR scattering rate is shown in Equation 2.15.

$$\Gamma_{\alpha\beta\kappa} = F\Gamma^0_{\alpha\beta\kappa} \tag{2.15}$$

where $\Gamma_{\alpha\beta\kappa}$ is the new scattering rate, $\Gamma^0_{\alpha\beta\kappa}$ is the scattering rate calculated using the previous method, and F is a dimensionless factor based off the ratio of L/ $\Lambda_\perp$ defined in Equation 2.16.

$$F = \frac{1}{V_o^2}\int dz_1 \int dz_2 \frac{\partial \overline{V}}{\partial z}(z_1)\frac{\partial \overline{V}}{\partial z}(z_2)C_\perp(|z_2 - z_1|) \tag{2.16}$$

However, Grange shows [9] that this can be simplified to a dimensionless parameter defined in Equation 2.17.

$$F = \exp\left(\frac{l^2}{16\,ln(2)\Lambda_\perp^2}\right)\left[\mathrm{erf}\left(-\frac{L}{4\,ln(2)\Lambda_\perp}\right) + 1\right] \tag{2.17}$$

How the incorporation of this updated version of IFR scattering affects lifetimes and leakage currents is discussed in-depth in Chapter 6; however, in general, the inclusion of the axial correlation length and interfacial mixing width leads to lower IFR scattering rates and lower IFR leakage for the same electronic temperature values.

References

[1] J. Faist, “Quantum Cascade Lasers”, Oxford University Press (2013)

[2] P. J. Price, “Two-dimensional electron transport in semiconductor layers i: phonon scattering,” Ann. Phys. (San Diego), vol. 133, no. 2, pp. 217-239, 1981.

[3] D. Botez, C.-C. Chang, L. J. Mawst, “Temperature sensitivity of the electro-optical characteristics for mid-infrared (λ = 3-16 µm)-emitting quantum cascade lasers,” *J. Phys. D: Appl. Phys*. vol. 49, no. 4 043001, 2016.

[4] R. Terazzi, “Transport in quantum cascade lasers”, Ph.D. dissertation, ETH Zurich, pp.103-106, 2011.

[5] C. Boyle, K. M. Oresick, J. D. Kirch, Y. V. Flores, L. J. Mawst, and D. Botez, “Carrier Leakage via Interface-Roughness Scattering Bridges Gap Between Theoretical and Experimental Internal Efficiencies of Quantum Cascade Lasers”, *Appl. Phys. Lett.,* **117**, 051101 (2020)

[6] A. Rajeev, C. Weixin, J. D. Kirch, S. E. Babcock, T. F. Kuech, T. Earles, and L. J. Mawst, “Interfacial Mixing Analysis for Strained Layer Superlattices by Atom Probe Tomography”, Crystals **2018**, 8, 437

[7] Y. Li, J. Zhang, W. Wan, Y. Zhang, Y. Nie, J. Zhang, and Y. Hao, “Alloy disorder scattering limited mobility of two-dimensional electron gas in the quaternary AlInGaN/GaN heterojunctions”, Physica E 67 (2015) 77-83

[8] T. Unuma, T. Takahashi, T. Noda, M. Yoshita, H. Sakaki, M. Baba, and H. Akiyama, “Effects of interface roughness and phonon scattering on intersubband absorption linewidth in a GaAs quantum well”, *Appl. Phys. Lett*. **78**, 3448 (2001).

[9] T. Unuma, M. Yoshita, T. Noda, H. Sakaki, and H. Akiyama, “Intersubband absorption linewidth in GaAs quantum wells due to scattering by interface roughness, phonons, alloy disorder, and impurities”, *J. Appl. Phys*. **93**, 1586 (2003).

[10] Y. T. Chiu, Y. Dikmelik, P. Q. Liu, N. L. Aung, J. B. Khurgin, and C. F. Gmachl, “Importance of interface roughness induced intersubband scattering in mid-infrared quantum cascade lasers,” *Appl*. *Phys. Lett*., **101**, 171117 (2012).

[11] Y. V. Flores, S. S. Kurlov, M. Elagin, M. P. Semtsiv, and W. T. Masselink, “S. S, Kurlov, M. Elagin, M P. Semtsiv and W.T. Masselink, “Leakage current in quantum-cascade lasers through interface roughness scattering,” *Appl. Phys. Lett*. **103**(16), 161102 (2013).

[12] T. Grange, S. Mukherjee, M. Montanari, L. Persichetti, L. Di Gaspare, S. Birner, A. Attiaoui, O. Moutanabbir, M. Virgilio, and M. De Seta, “Atomic-Scale Insights into Semiconductor Heterostructures: From Experimental Three-Dimensional Analysis of the Interface to a Generalized Theory of Interfacial Roughness Scattering”, *Phys*. *Rev. Appl.*, **13**, 044062 (2020)

[13] S. Mukherjee, A. Attiaoui, M. Bauer, O. Moutanabbir, “3D Atomic Mapping of Interfacial Roughness and Its Spatial Correlation Length in Sub-10 nm Superlattices”, *Appl. Mater. Interfaces* 2020, **12**, 1728-1739

CHAPTER THREE

FAILURE ANALYSIS OF QCLs

3.1 Introduction

Failure mechanisms for interband-transition semiconductor lasers are well studied and supply a solid roadmap for the identification and mitigation of failure modes for intraband transitions. In the case of visible and near-infrared diode lasers, the main point of failure occurs at the front facet and is typically characterized by sudden catastrophic optical damage (COD), sometimes referred to as catastrophic optical mirror damage (COMD) [1]. The front and back mirrors are cleaved such that they have two parallel facets that act as the mirrors of a resonator. However, the mirrors are dissimilar to the bulk of the laser, due to dangling bonds at the surface. During device operation there is strong preferential absorption at the front facet, as typically the back facet has a high reflectivity (HR) coating, thus less light intensity that at the front fact, as well as better heat removal, which leads to strong heating at the front facet. As the temperature at the front facet increases, there is a narrowing of the band gap and an increase in the surface recombination rate [2]. Elevated temperatures and surface recombination rates are in a positive feedback loop and can lead to facet degradation and/or COMD [1].

The use of intraband transitions circumnavigates the narrowing of the band gap at the front facet as the band gap is no longer utilized. Previous lifetest studies of QCLs [3] have derived an activation energy of 1.2 eV which translates to a mean time to failure (MTTF) of 809,000 hours. Although that study was performed at relatively low output powers (<0.2W) it did show the potential for stability and reliability of QCLs.

A reliability rack was constructed to begin lifetest studies and failure analysis of high-power (0.5-1.0 W) QCLs operating CW at room temperature. This rack was outfitted with 19 channels that can run individually monitored and controlled devices in a few different configurations. For the work presented here, all the devices were run under constant power where the device current is monitored and adjusted to maintain constant power. however, the ability to run under constant current or constant voltage was also present. To have the ability to run constant power, the rack included 19 separate current drivers, thermo-electric coolers (TEC's), Mercury-Cadmium-Telluride (MCT) detectors supplied by Vigo Systems, 2" gold integrating spheres, and water-cooling blocks. Each of these channels were actively cooled with the TEC and individually driven with a Thorlabs ITC4005 current driver. The devices were in a closed particle-controlled environment with the inlet filtered with a high-efficiency particulate air (HEPA) filter to maintain a clean environment. The temperature of the devices was monitored using a PT100 which was placed approximately 1 mm to the side of the chip. As will be discussed in Section 3.2.2, within the range of 0.5 mm to 1.5 mm, the positioning of the temperature sensor is relatively insensitive as the Au/Cu submount has an extremely high thermal conductivity.

The initial plan was to separate the devices into two different groups; one would be run with the submount temperature held at 20 °C, while the other would be at an elevated temperature. However, as the first group at 20 °C was running, multiple failure mechanisms outside of the active region emerged that needed to be addressed. Similarly, at elevated temperatures, failures outside of the active region became apparent. The following sections discuss the failure mechanisms that were observed and what steps were instituted to mitigate these failures and/or what steps could be taken to further their reliability.

3.2 Identified Failure Mechanisms

Devices used in the lifetest rack were 40-stage strained step-tapered active-resonant extraction (STA-RE) [4] buried heterostructure (BH) material emitting at ~5.0 µm with the following CW light-current-voltage (LIV) and wall-plug efficiency (WPE).

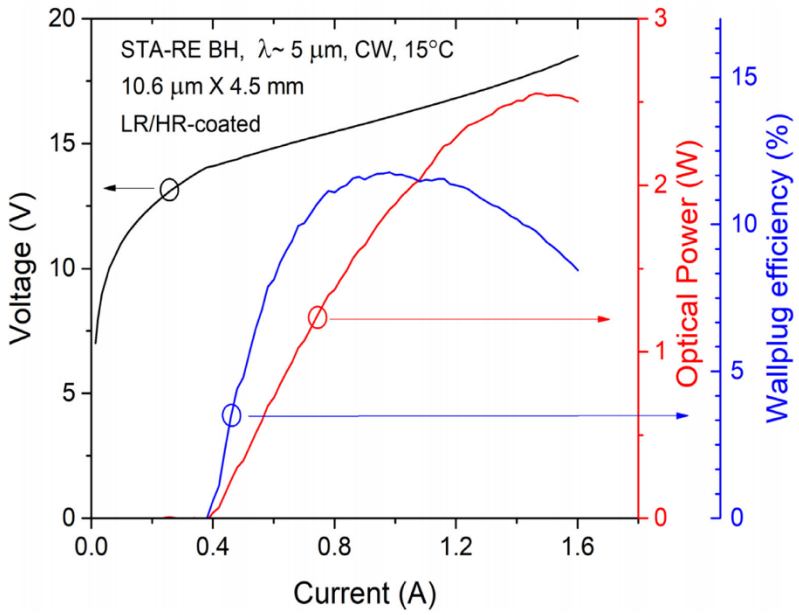


Fig. 3.1. LIV for a 10.6 µm wide device mounted epi-down on copper with indium under CW operation [4].

A scanning electron microscopy (SEM) image of the cross-section of a typical device from this run can be seen in Figure 3.2 (a) below. The calibration of these 10 µm x 5 mm devices was performed at 10% duty cycle (1 kHz and 1 µs pulse width) and the reproducibility across multiple devices and bars can be seen in Figure 3.2 (b). At this duty cycle, at the device level, QCW operation is achieved, however, the thermal dissipation requirements is far more manageable than in CW operation. Devices were calibrated at 10% duty cycle using a thermopile which was then correlated with the MCT response for each specific channel and device. Once the correlation factor was ascertained, the output power, being read in A.U., could be correlated with a calibrated power measurement. Devices were operated at constant power by

monitoring this output signal and holding the target value which corresponded to a calibrated power of 1 W CW.

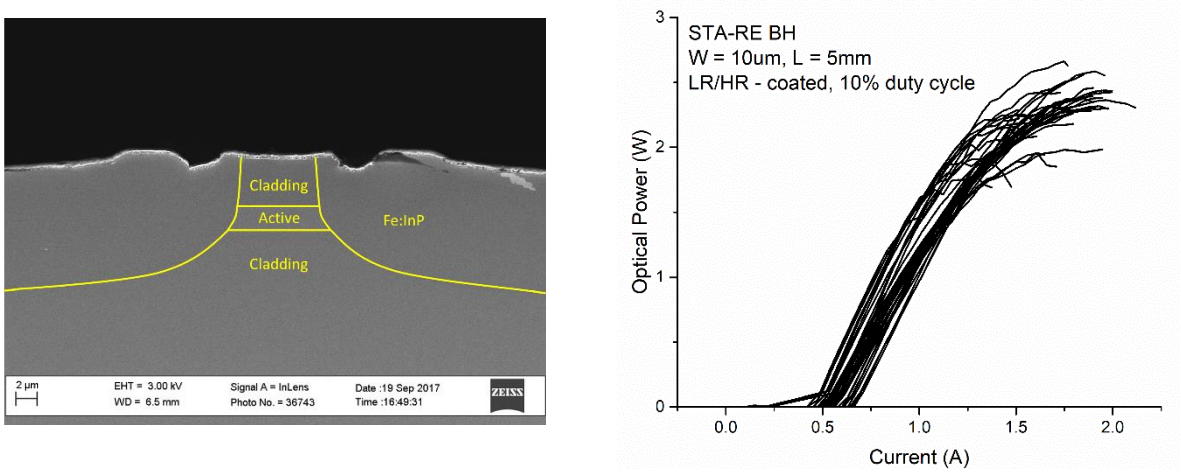


Fig. 3.2 (a) SEM image showing the cross-section of a BH device from the same material to be mounted for the lifetest rack, with the yellow lines as a guide to show the respective device regions, and **(b)** the calibrated power for devices to be mounted in the reliability test rack.

To mitigate the failure observed by Xie et al. [3], a low reflectivity (LR) coating was applied to the front facets, acting as both a LR coating and a passivation layer. To achieve high-single facet power, HR coatings were applied to the back facets. The devices were then run at 1.0 W CW output power with submount temperatures of 20 °C and 50 °C. Under these conditions multiple failure mechanisms were identified and mitigated.

3.2.1 Catastrophic Mirror Damage

Like diode lasers, catastrophic mirror damage is observed, but the instigation of this failure is not well understood. Dissimilar to diode lasers, QCLs have a much higher thermal management requirement. At a wall-plug efficiency of 10% and using the LIV shown in Figure 3.1, the heat dissipation requirement is about 25W compared to about 1 W heat dissipation required for a 940 nm diode. While 25 W heat

removal may seem trivial, this is being applied to an extremely small area, around $4.8 * 10^{-4}$ cm^2. This results in a large thermal gradient between the active region and the substrate which is believed to lead to this failure mechanism [5, 6]. Catastrophic mirror damage from these studies, as shown in Figure 3.5 (b), exhibits a nearly identical semi-circular shape as found by Zhang [5] (see Figure 3.3 (a)).

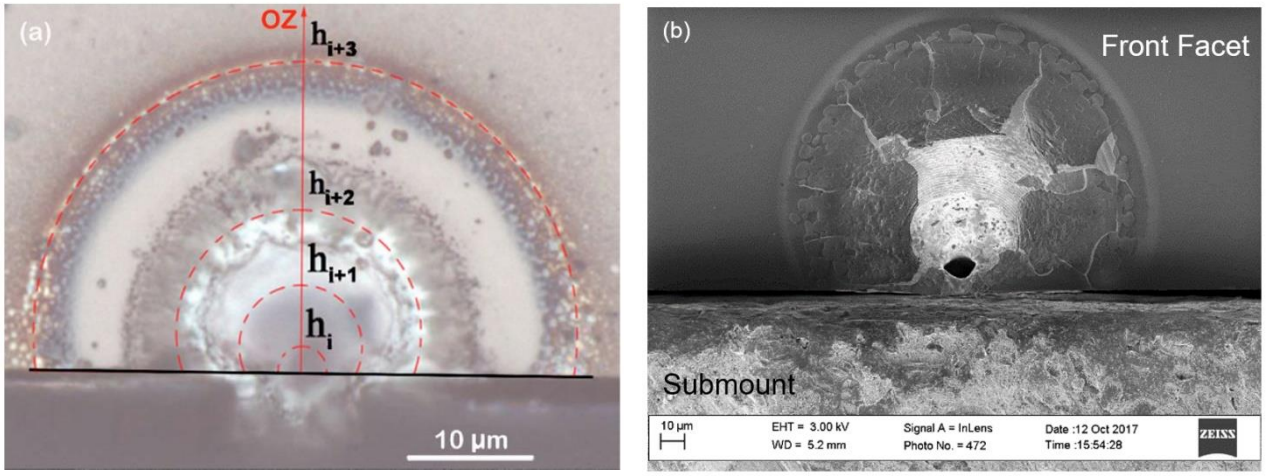


Fig. 3.3 (a) Failed ridge guide structure from Zhang which shows the semi-circles believed to be stress relaxation lines [5] **(b)** QCL device from the lifetest study mounted epi-down on copper with indium which failed after control of the TEC was lost and the temperature rapidly increased.

This type of failure was mostly instigated by the loss of control of the TEC. With such large heat removal requirements over such a small area the limits of TEC technology to maintain 20 °C on the submount is being pressed. Under high loads the TEC can no longer maintain the desired temperature, or the voltage/current limit of the TEC is reached, and this failure mechanism was observed. Through optimization of the software the amount of catastrophic mirror damage failures is significantly reduced by immediately removing power from the device once a predetermined variation in temperature is observed. This failure was also mitigated by reducing the thermal resistance of the fixturing using PGS graphite sheets for lateral heat transport, gold metallization, polishing, and reducing the acceptable

temperature fluctuations of the device during lifetest operation. This window was reduced to ± 0.5 °C such that if there was a temperature swing that brought the device out of this specification, the device would immediately turn off. Ideally, this would save the device from this type of failure mechanism, however, while this fail-safe measure did save the majority of the devices, in some cases the reaction of the current driver to turn off was too slow and the device still died. Work is being performed to reduce the lag time between the temperature flag and the current source being turned off.

With the help of Dr. Yongkun Sin from The Aerospace Corporation this mode of failure was investigated with transmission electron microscopy (TEM) and as a function of position with respect to the front facet via focused-ion-beam (FIB) cross-sections. To remove the device from the copper submount in order to use FIB, one device was warmed to the point that it could be removed from the indium (In) soft solder. Devices that were mounted epi-down with AuSn solder could not be removed. FIB was used to track the failure back into the device to see if a defect or place of origin could be determined. The device had been mounted epi-down on copper with In and was operated at a submount temperature, measured 0.7 mm from the device, of 22.5 °C +/- 0.5 °C. The device also had 100 nm of evaporated yttria for a LR coating. After nearly 15 days of 1 W CW constant power, the device failed. The current as a function of time can be seen in Figure 3.8 below.

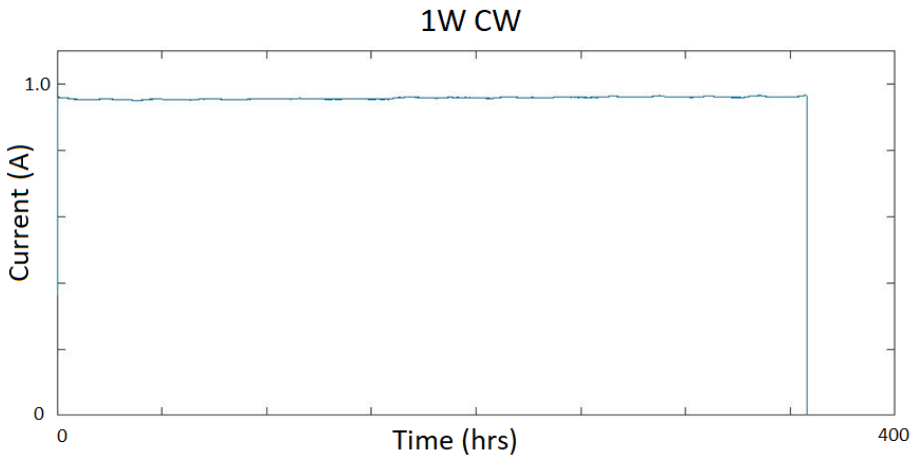


Fig 3.4 Lifetest data for device BF90 which failed after nearly 15 days of 1 W CW constant power at a submount temperature of 22.5 °C and was investigated with cross-section FIB and TEM.

Looking at this device with an optical microscope, failure directly over the active region can be seen. Optically, it is difficult to see (see Figure 3.5), but under SEM, shown in Figure 3.6, the typical catastrophic mirror damage with semi-circles radiating away from the active region can be clearly seen.

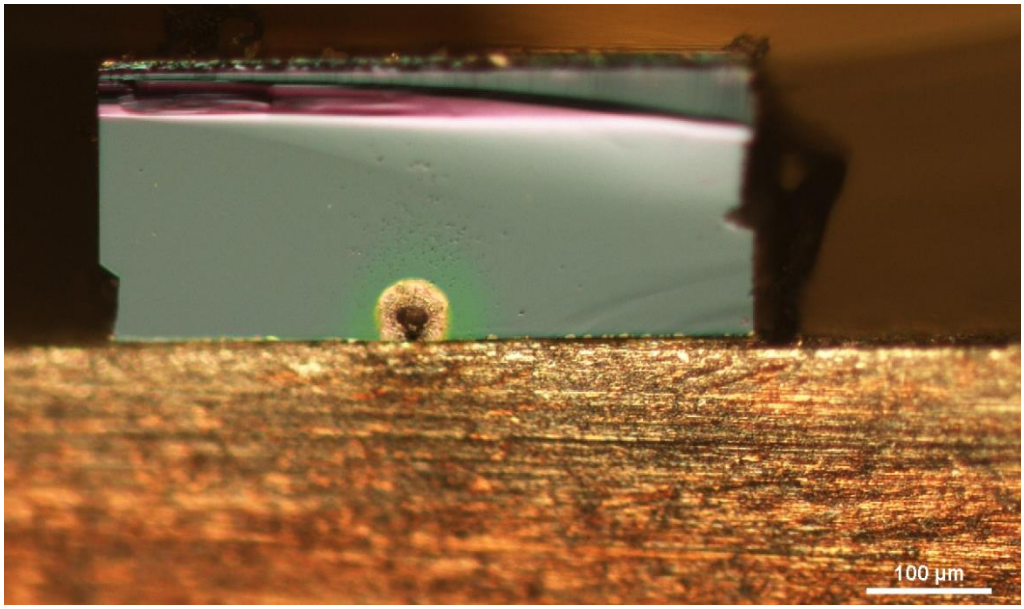


Fig 3.5 Optical microscopy image of device BF90's front facet after failure. The damaged region is directly over the active region of the device.

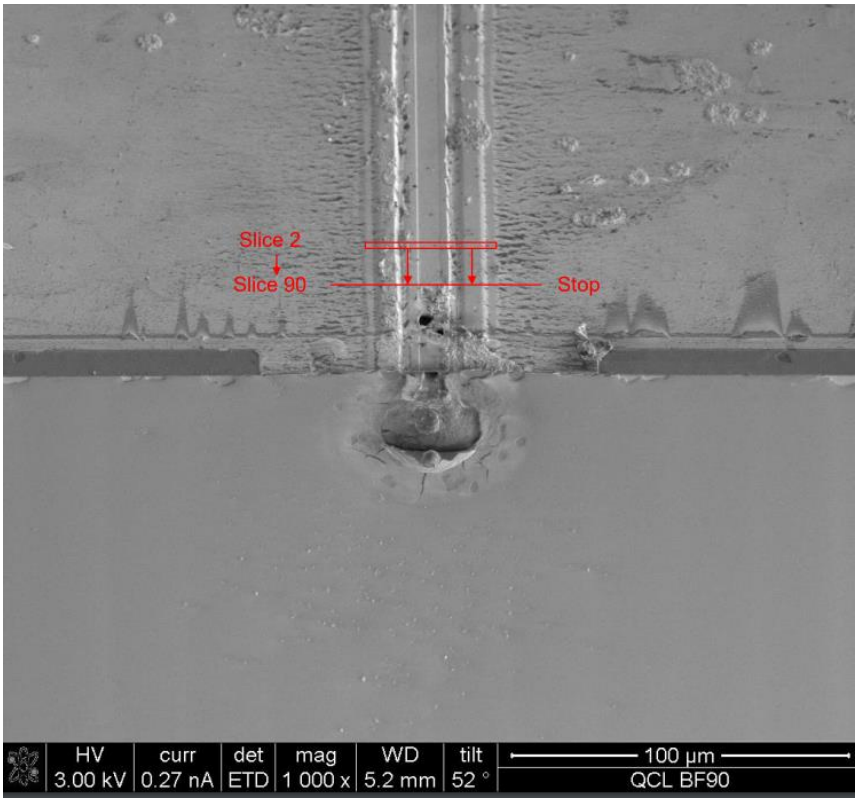
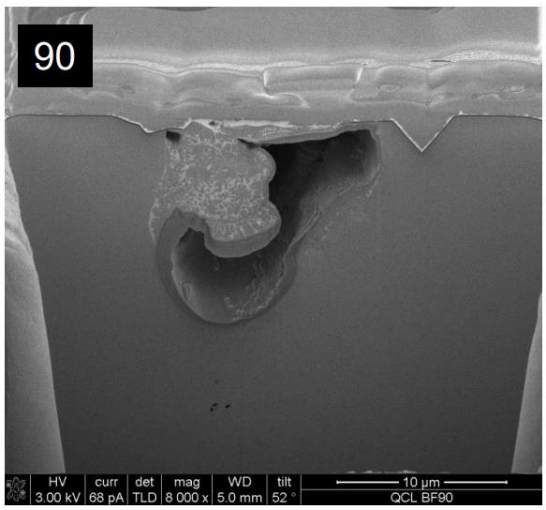
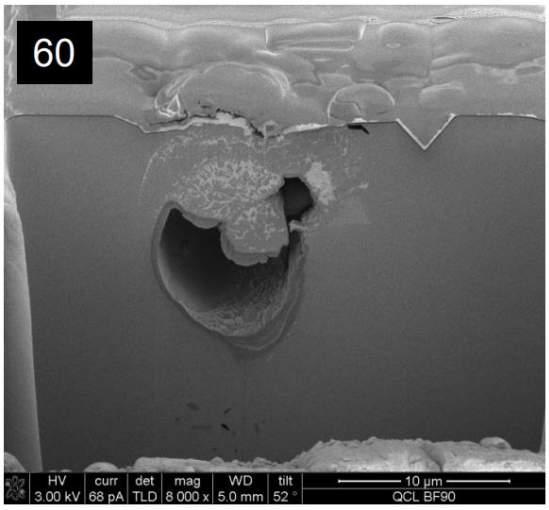
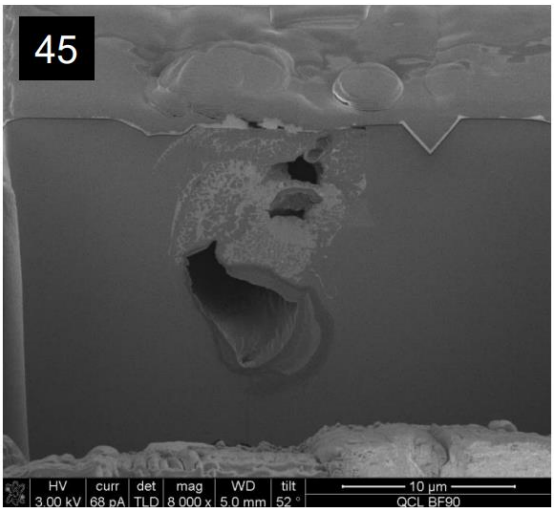


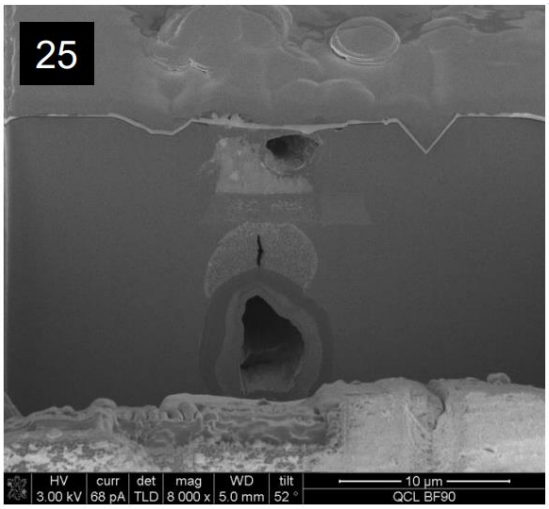
Fig 3.6 SEM image of the BH device that was previously mounted epi-down on copper with indium, experienced failure, and shows catastrophic mirror damage directly over the active region.

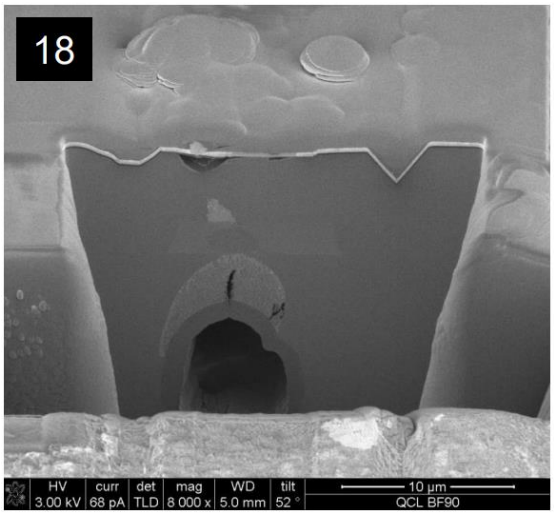
The SEMs below show the device cross-section as a function of position moving away from the front facet towards the center of the device.

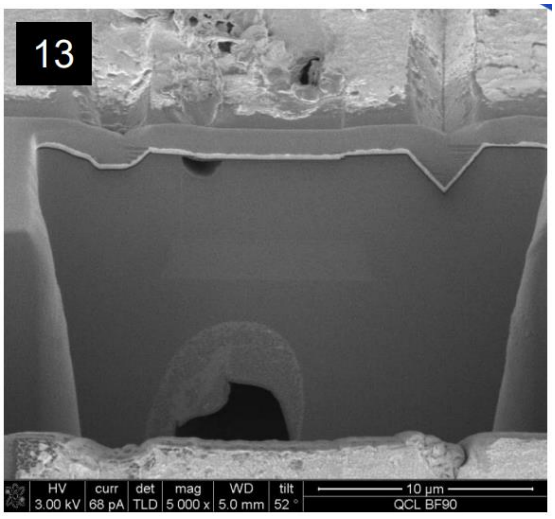

90
HV curr det mag WD tilt
3.00 kV 68 pA TLD 8 000 x 5.0 mm 52 °
10 µm
QCL BF90


60
HV curr det mag WD tilt
3.00 kV 68 pA TLD 8 000 x 5.0 mm 52 °
10 µm
QCL BF90


45
HV curr det mag WD tilt
3.00 kV 68 pA TLD 8 000 x 5.0 mm 52 °
10 µm
QCL BF90


25
HV curr det mag WD tilt
3.00 kV 68 pA TLD 8 000 x 5.0 mm 52 °
10 µm
QCL BF90


18
HV curr det mag WD tilt
3.00 kV 68 pA TLD 8 000 x 5.0 mm 52 °
10 µm
QCL BF90


13
HV curr det mag WD tilt
3.00 kV 68 pA TLD 5 000 x 5.0 mm 52 °
10 µm
QCL BF90

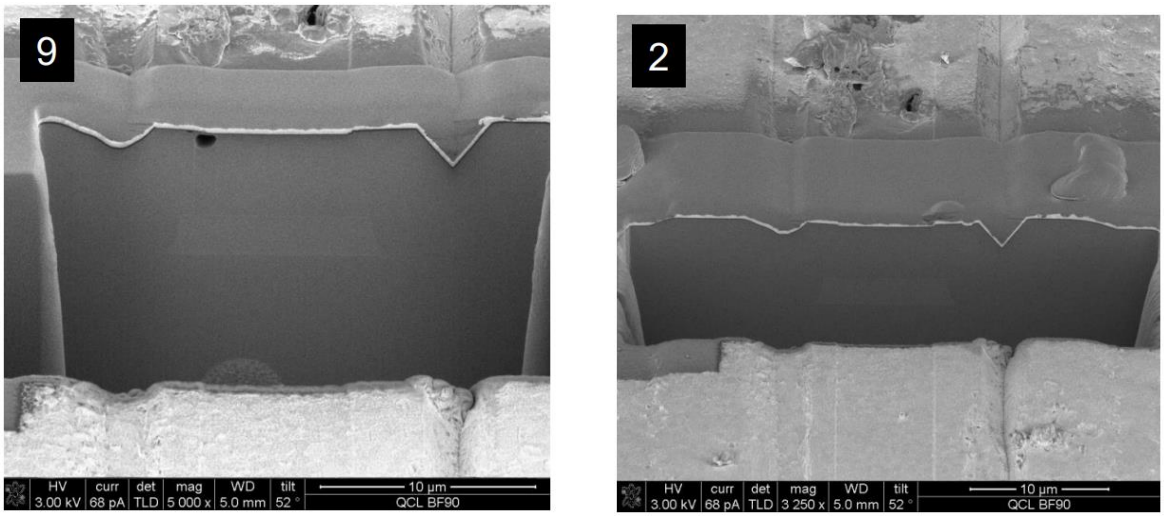


Fig 3.7 As the numbers decrease, the cross-section is moving further into the cavity and away from the front facet.

Tracking the voids back into the cavity it looks like they are originating both towards the substrate, where the greatest thermal stress is expected, as well as near the metal contact. It has been shown previously [7] that, under high current densities for extended periods of time, formation of voids at the platinum/gold interface can occur.

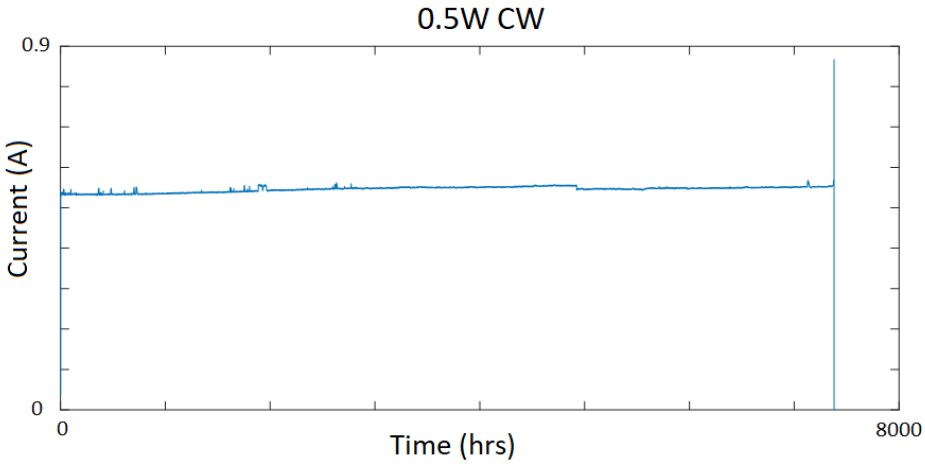


Fig. 3.8 Lifetest data for a device running at a constant power of 0.5W CW with a submount temperature held at 20 °C for nearly 300 days.

This FIB process was then repeated on a device that similarly failed, but no longer with a front-facet coating. This device was operated at 0.5 W CW constant power with the submount temperature maintained at 20 °C. After nearly 300 days of continuous operation, the lifetest graph can be found in figure 3.8, this device also experienced a catastrophic mirror damage. As the voids are tracked back into the cavity, as shown in figure 3.9 below, track to locations above and below the active region similar to the previous device.

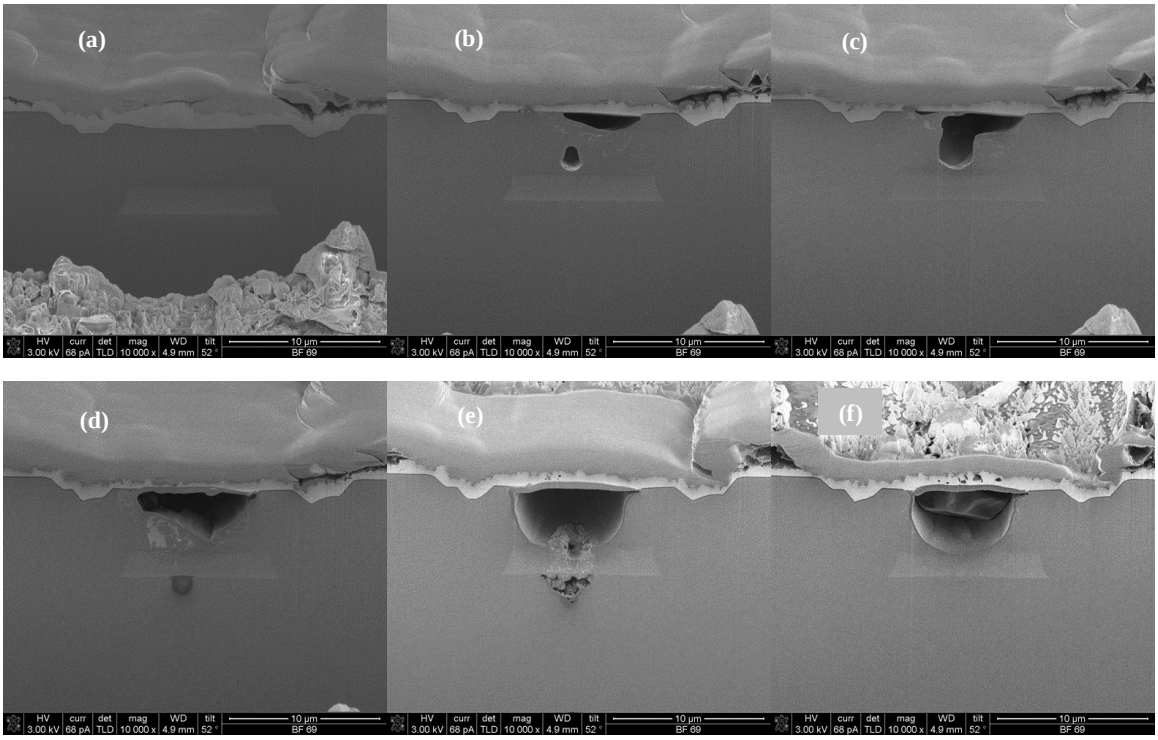


Fig 3.9 SEM cross-sections of device BF16b starting in the undamaged cavity in **(a)** and moving towards the front facet moving from **(a-f)**.

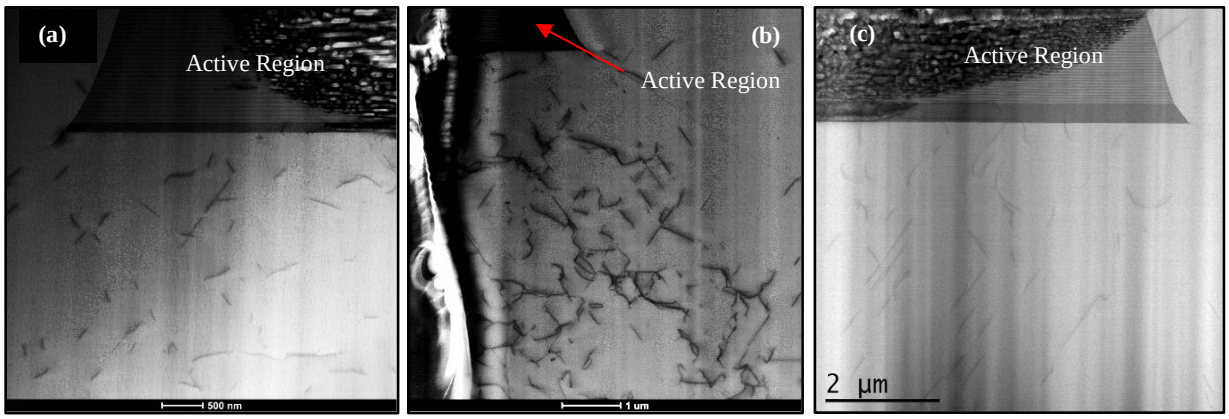


Fig 3.10 TEM images performed at the Aerospace Corp. showing threading discolations below the active region of BF16b after failure.

In quantum wells (QW), COMD typically shows threading dislocations within and around the well; however, these images show the dislocations below the active region, and a polycrystalline region within the active suggesting potential localized melting [8].

Through proper heat management and bonding techniques, this mode of failure can be largely avoided as it is believed to be thermally instigated. This can be achieved through a few different approaches such as: optimizing the packaging to lower the thermal resistance; running the devices under QCW operation to lower heat generation which is addressed in the following section; through waveguide-geometry optimization which will be addressed in Chapter 4 and increasing the wallplug efficiency through IFR engineering which is discussed in Chapter 5; and through the characterization and optimization of interfaces which is discussed in Chapter 6.

3.2.2 Packaging

The close monitoring, control, and minimization of the temperature of QCLs are of the utmost importance for reliable operation. To accurately monitor the temperature of the active region a PT100 thermistor is placed < 1 mm from the device.

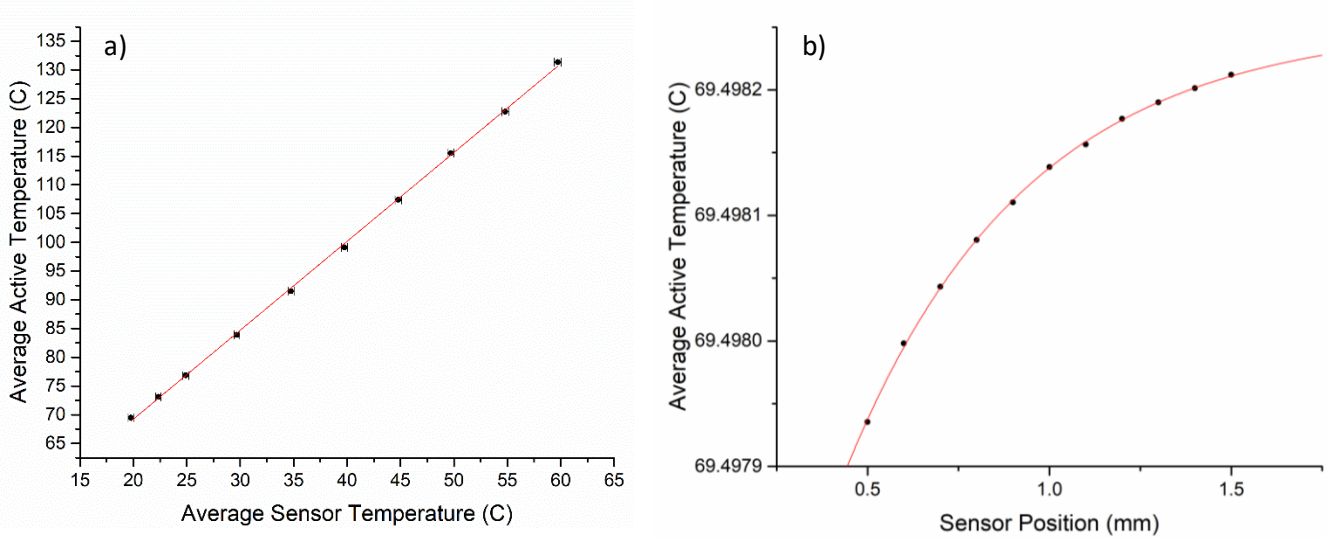


Fig 3.11 The average active-region temperature **(a)** calculated in COMSOL as a function of the temperature sensor reading place (1 mm away) assuming a 16 µm x 5 mm device mounted epi-down on copper with indium, and 20W of input power; and **(b)** assuming the thermistor reads a constant value of 20 °C and varies the position of the thermistor while maintaining a measured temperature of 20 °C.

It is clear from Fig 3.11 (a) that the active region is significantly hotter than the measured temperature and from Fig 3.11 (b) the placement of the thermistor with respect to the device is not hypersensitive. An accurate model for the active-region temperature with respect to the measured submount temperature is essential for furthering the investigation of reliable packaging of QCLs.

Initially, devices were mounted epi-down on copper with indium. This is a method that lends itself well to CW high-power operation as the soft solder, indium, can absorb the difference in thermal expansion coefficients between itself, the InP-based device, and the copper submount. Unfortunately, indium is known to creep at long time scales as well as under pulsed operation and is not appropriate for high current-density operation [9]. Below in Fig 3.12 are the results for 6 devices mounted epi-down on copper with indium and run under constant power at 1 W CW with the submount temperature of 50 °C. These devices quickly resulted in shorts caused by indium migration instigated by the elevated temperatures.

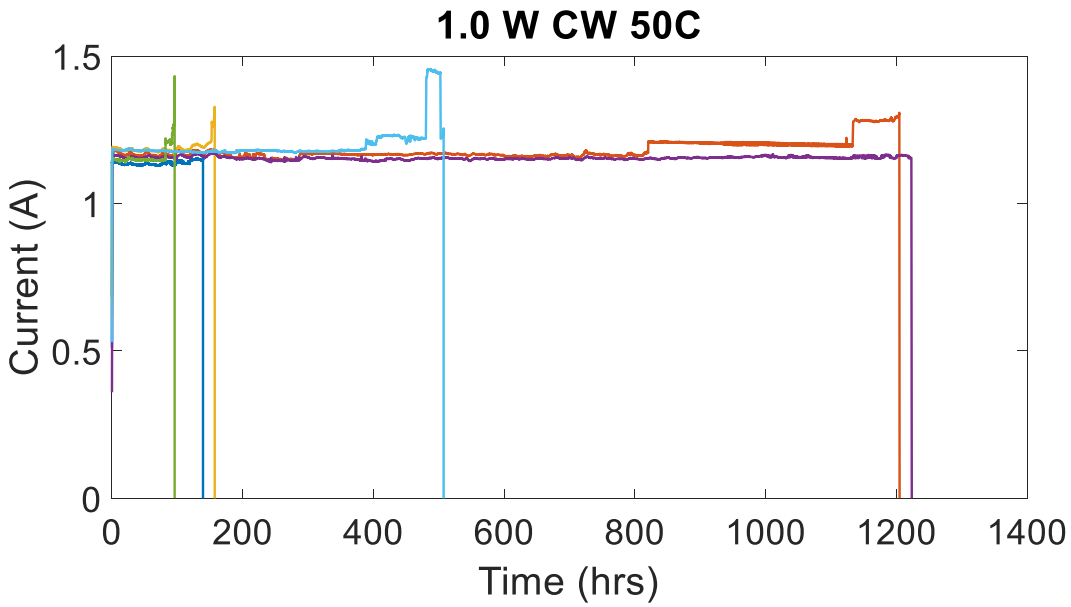


Fig 3.12 Lifetest results for 6 devices mounted epi-down on copper with indium running at an elevated submount temperature (50° C) running at 1 W CW power (under constant-power operation). All devices quickly failed as a result of indium migration.

To move towards the extraction of activation energies and run long-term reliability studies, a transition away from indium was necessary. Since there is no longer a soft solder to absorb discrepancies in the coefficients of thermal expansion (CTEs), a move to a fully CTE-matched package was required.

Table 1.1 List of semiconductor materials, submount materials, bonding materials, and their thermal conductivity and coefficient of thermal expansion relative to InP. ** $Au_{0.71}Sn_{0.29}$ is often referred to as AuSn (80/20). $Au_{0.71}Sn_{0.29}$ At% is AuSn (80/20) by weight.

Material	Thermal Conductivity (W/mK)	CTE (ppm/C)	CTE % Change with respect to InP
InP	68	4.5	-
SiN	10-43	1.4-3.7	-69% to -18%
Copper	390	17	278%
AlN	170-200	4.5	0%
SiC	200-300	3.7	-18%
$Cu_{0.1}W_{0.9}$	130	5.6	24%
$Cu_{0.25}W_{0.75}$	220	9.0	100%
$Au_{0.71}Sn_{0.29}$ **	57	16	256%
SAC305	58.7	23.5	422%
Indium	86	29	544%

Previous reliability studies have been performed on CuW with AuSn [3] and despite the relatively large CTE mismatch between InP and CuW, there were no reported failures. AuSn is a highly desired bonding medium as it does not require flux, has great wetting capabilities, and is stable long-term up to temperatures more than 300 °C. The phase diagram for AuSn is shown below in Fig 3.13.

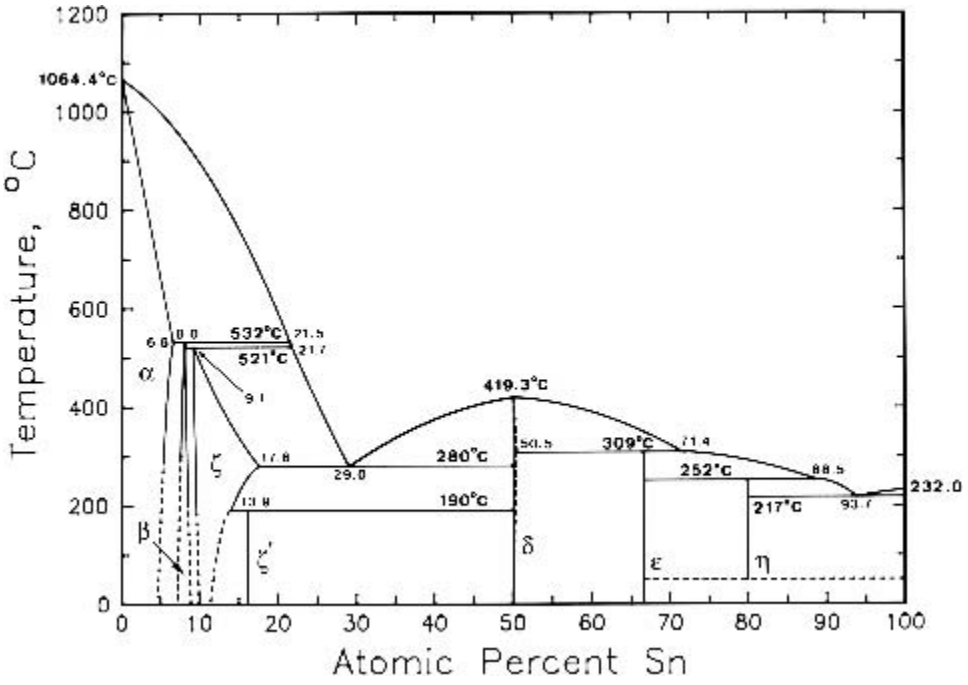


Fig 3.13 Phase diagram for the binary AuSn displaying a eutectic point at 29 At% Sn.

At an atomic percent Sn of 29% this solder exploits its, relative to the higher concentration Au side of the eutectic, low temperature eutectic point. Once the reflow occurs, diffusion from the two gold interfaces on either side diffuse in reducing the Sn atomic percentage. Once the bond cools and reaches its solidus temperature, the liquidus temperature is shifted up the curve to the left to reflect the lowered atomic percent Sn resulting in significantly higher liquidus temperatures. This is typically indicative of lowered creep.

The active-region temperature can now be monitored and the relation between submount temperature and active-region temperature has been found. The ability to mount devices on CuW and/or

AlN with AuSn has also been incorporated to circumnavigate the indium migration that was previously resulting in failures under QCW operation and at elevated temperatures.

3.2.3 Low-Reflectivity Coating

As mentioned previously, a low-reflecticity (LR) coating was applied to the front facet to avoid oxidation at elevated temperatures. The coating is 100 nm of electron-beam evaporated yttria with a silicon adhesive layer. This coating was applied while the devices were still in bar form so that it was applied uniformly across all devices, however, this implies the coating is deposited prior to device mounting. Mounting at temperatures required to reflow AuSn (~320 °C) this LR coating has shown delamination.

The devices studied here were mounted with indium solder so the reflow temperature is much lower than required by AuSn, so the failure mentioned is not observed for these specific devices. However, moving forward the coatings will need to be deposited after the devices are mounted, and/or develop LR coatings that can be deposited with sputtering for more dense films that may be more resilient to the high temperatures that are necessary for AuSn bonding.

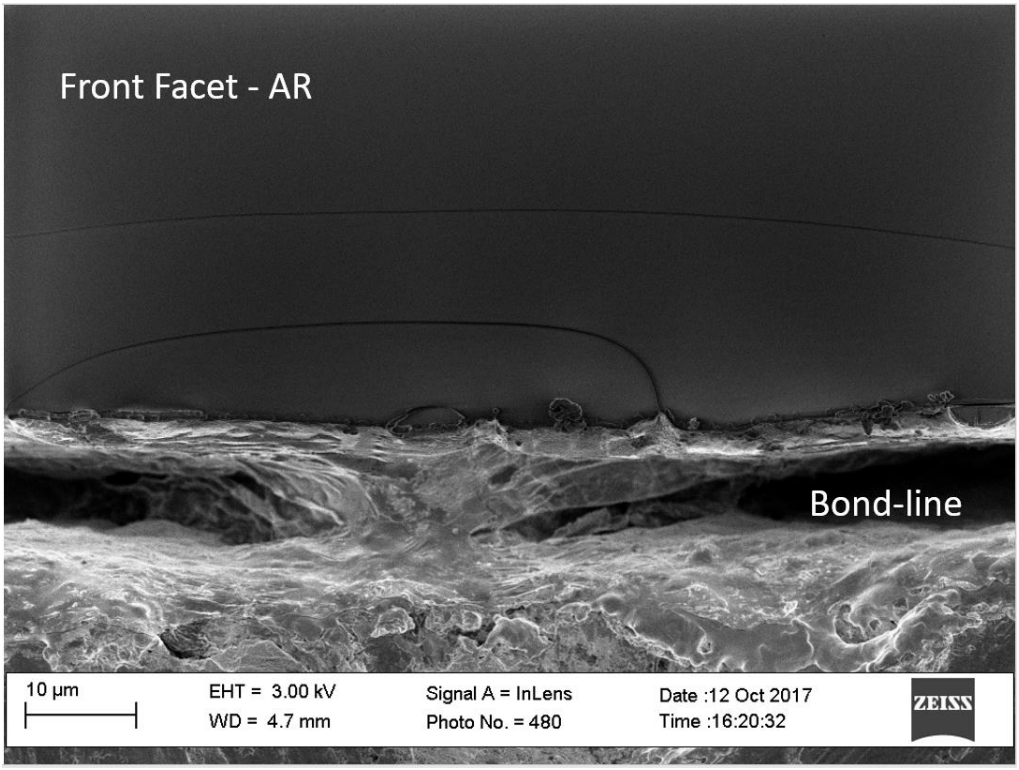


Fig. 3.14 SEM image of the front facet of a QCL device with a 100 nm thick yttria anti-reflective (AR) coating after being mounted on CuW with AuSn.

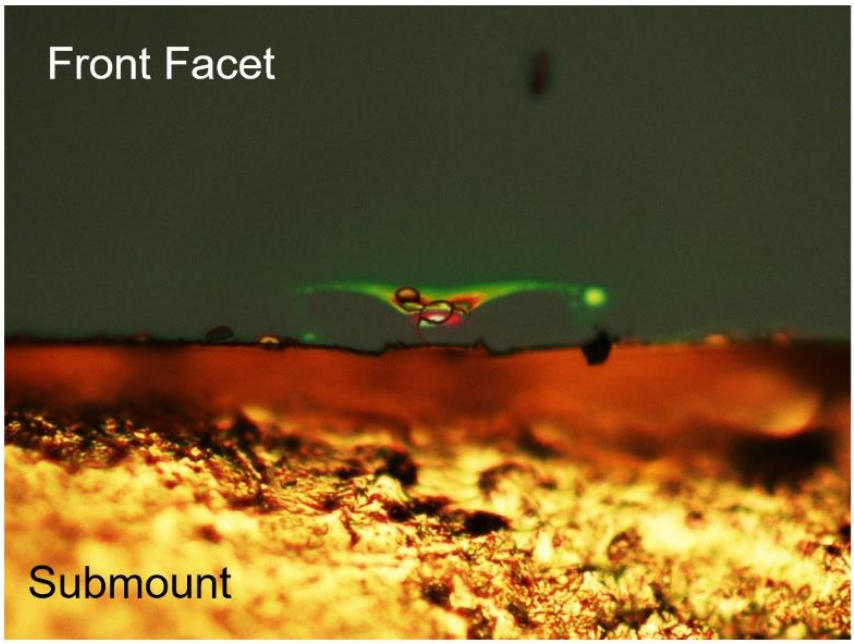


Fig. 3.15 Optical image of the front facet of a QCL with a 100 nm thick yttria AR coating mounted on Cu with In after running for 2000 hours at 1W CW.

Despite the lower temperature of the indium bonding, failures in the LR coating as a slow degradation of the output power are still observed. An optical image of the degradation in the LR coating is shown in figure 3.15 and the decrease in power and slope efficiency are shown below in figure 3.16.

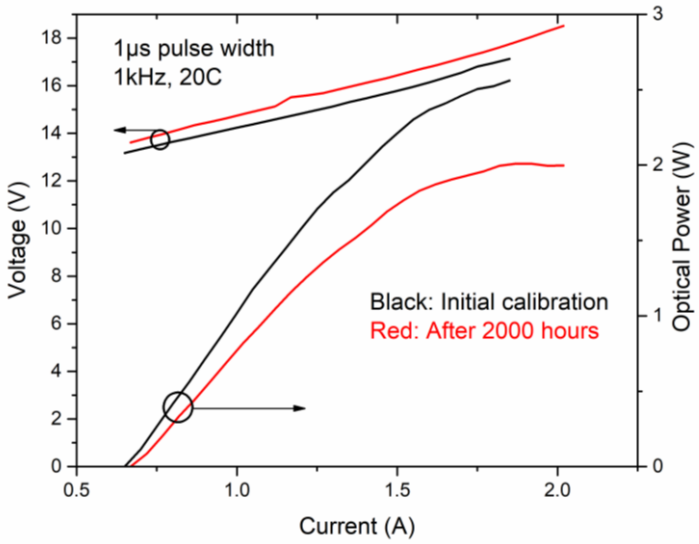


Fig 3.16 Calibrated LIV for a device under 10% duty cycle measured with a thermopile at a temperature of 20 °C before and after the device was run for 2000 hours at 1W CW with a submount temperature of 20 °C.

A potential route explored to circumvent the need for evaporated or sputtered LR coatings was through the use of milled facets as shown by Dirisu et al [10]. Ridge guides were mounted epi-up with conductive silver epoxy on copper submounts. These devices were tested, and the LIV was gathered before milling the front facets. The figures below show the failure directly over the active region for devices with two different periodicities of the grating.

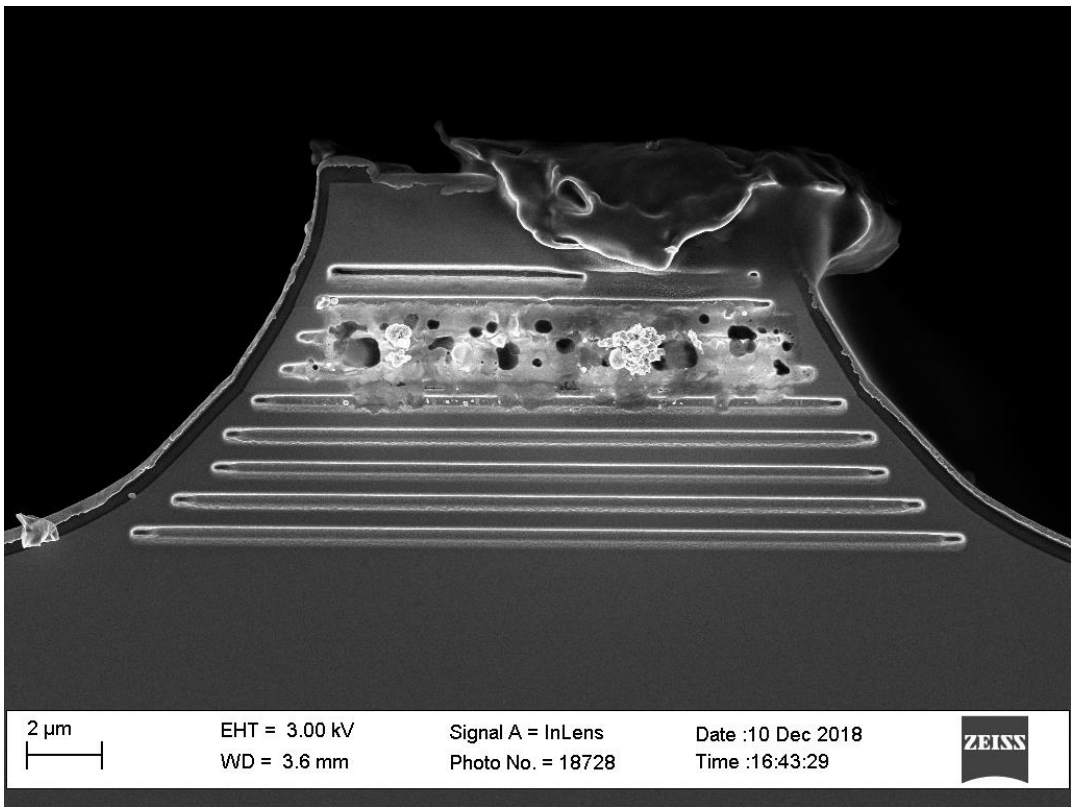


Fig 3.17 SEM image of a failed ridge guide with a milled LR coating.

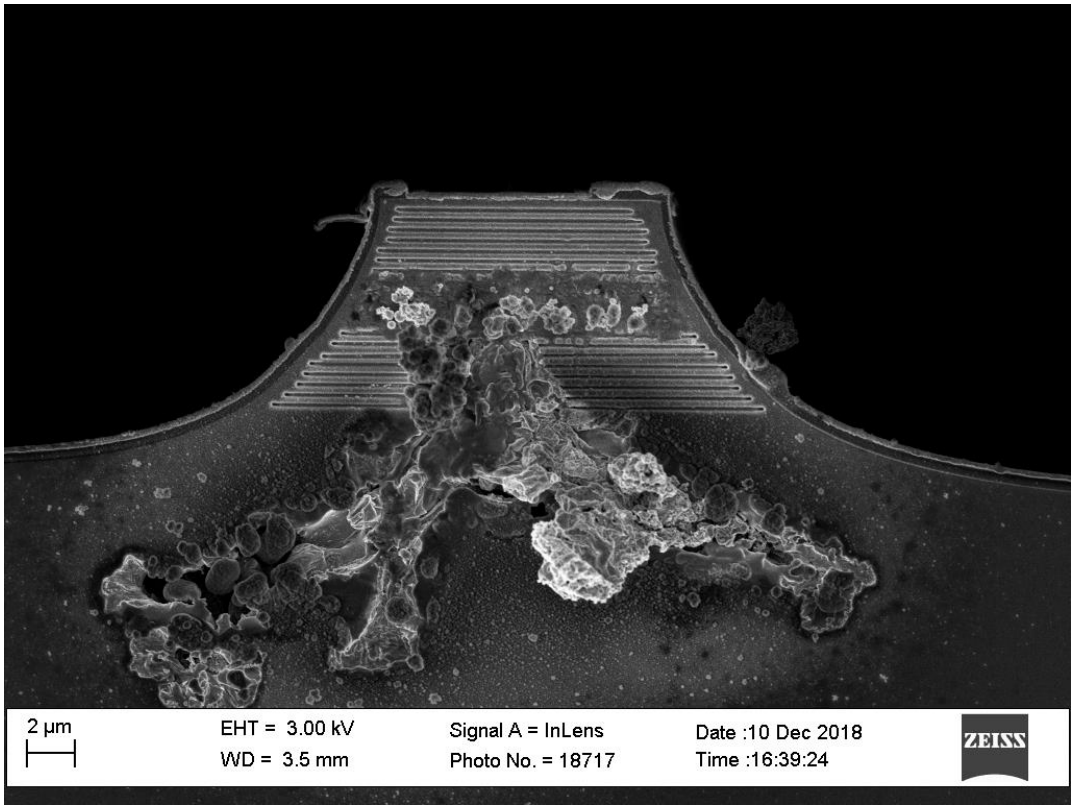


Fig 3.18 SEM image of a failed ridge guide with a milled LR coating.

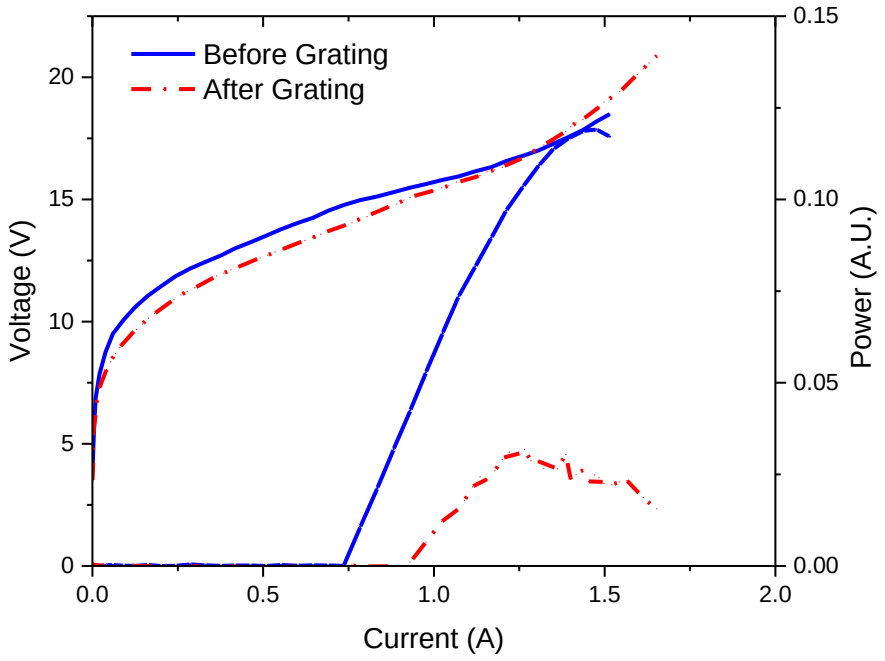


Fig. 3.19 LIV of the device shown in Fig. 3.17 before and after etching the grating in the front facet.

It can be seen from Figure 3.19, while these devices did exhibit higher thresholds and lower slopes as one would expect from a device with an AR or LR coating, they quickly died. The FIB work was done with a Ga source. This would most likely lead to gallium implantation in the front facet during the milling. The Ga implantation could act as a soft short for this device which would explain the lower voltages below threshold. This may also explain the failure at the front facet that is being observed. While this is an interesting option for lowering the front facet effective reflectivity, the reliability of this method is suspect. A potential route forward for using this method is to use a FIB that has an inert source, such as Xenon. However, for the meantime, devices without a front facet coating appear to be the most reliable and repeatable option.

3.2.4 High-Reflectivity Coating

One of the first observed failure mechanisms was issues with the high-reflectivity (HR) coating. Under the high thermal loads experienced by the device, the mismatch in thermal expansion coefficients lead to flaking and delamination of the back facet as exhibited in the figure below.

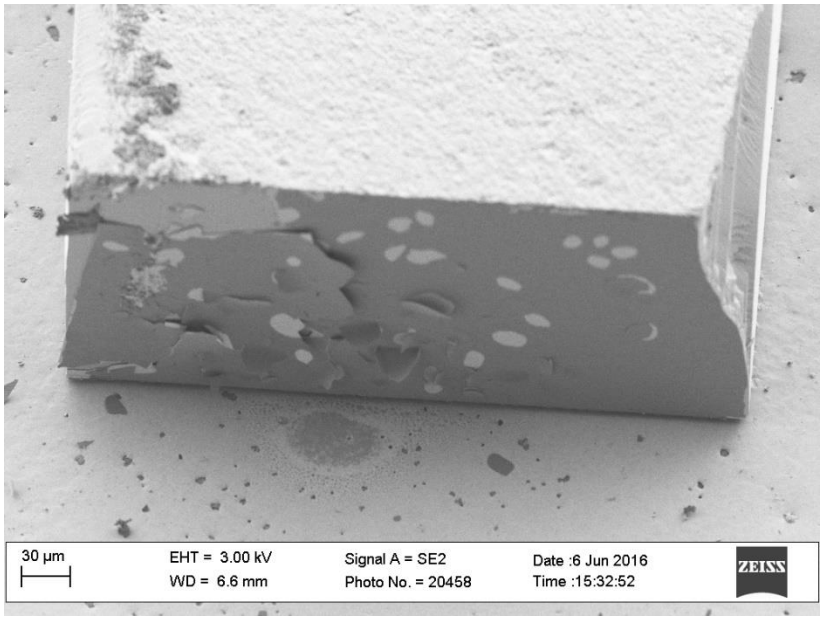


Fig 3.20 Back facet of a QCL mounted epi-down on copper with indium showing the HR coating is flaking off due to the high thermal stresses induced during CW operation.

To address this concern an athermal back mirror was developed. Using a combination of adhesion layers, dielectrics, and metal with varying thicknesses the CTE of the film strain-balanced to InP. This CTE matched to InP film is able to consistently sustain CW and QCW operation. The following studies are all performed using this updated CTE-matched back mirror.

As previous reliability studies had been performed on CuW with AuSn; this was the first route explored despite the CTE mismatch described previously. Another failure at the back facet was observed with the move from indium-mounted devices to AuSn-mounted devices. This failure mechanism is shown in the figure below.

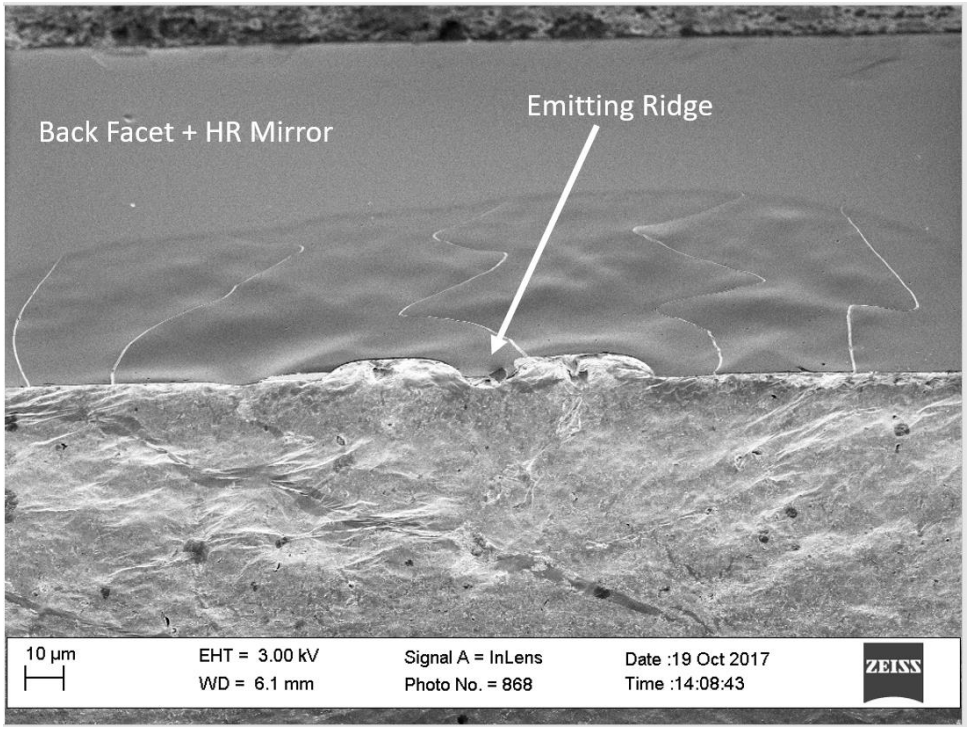


Fig. 3.21 Back facet of a ridgeguide QCL mounted epi-down on AuSn showing the HR mirror failure during the AuSn bonding process.

This failure is observed before the device is biased and it is believed to be caused by the gold diffusion from the AuSn solder into the gold in the back mirror. To block this diffusion, 100nm of Al_2O_3 overspray is used to cap the back mirror on the edge where the HR coating meets the top metal and where the AuSn has the opportunity to diffuse into the metal in the back mirror. The figure below shows the fixturing fabricated for this process.

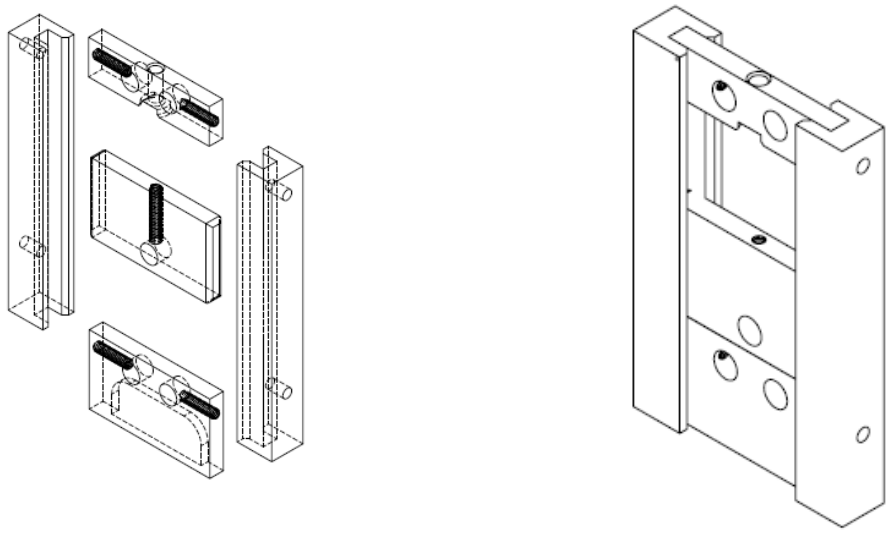


Fig. 3.22 Computer-aided drawing of a redesigned bar holding fixturing to deposit a dielectric overspray at the interface between the back facet HR and the epi-side of the bar.

When using this fixturing, the bars are aligned to one edge and then stacked on top of each other with double-polished silicon spacers in between. This fixture is quite thin (~5 mm) which allows for deposition on both facets. The extent of the alumina overspray onto the epi-side can be controlled by scribing the silicon spacer bars to different lengths. Devices have been successfully mounted with AuSn with an overspray of 100 µm and 50 µm. This 100 nm of alumina overspray, while acting as a diffusion barrier, also is an additional thermal barrier and as such the thickness should be minimized. The back facets, after addressing the facet coating concerns, has not been an area of concern regarding failure. Device failures are believed to be thermally instigated, and through COMSOL Multiphysics it has been

observed that the front facet of epi-down devices are hotter than the back facet. This is caused by two different mechanisms: 1) the back facet coating helps remove heat from directly over the active relative to an uncoated front facet, and 2) since the front facet is aligned to the edge of the submount, there is a direction of heat removal that is absent relative to the back facet.

Devices were successfully mounted epi-down on CuW with AuSn and alumina overspray. These devices were then put on life-test and run until failure. The life test for these two devices is shown in the figure below.

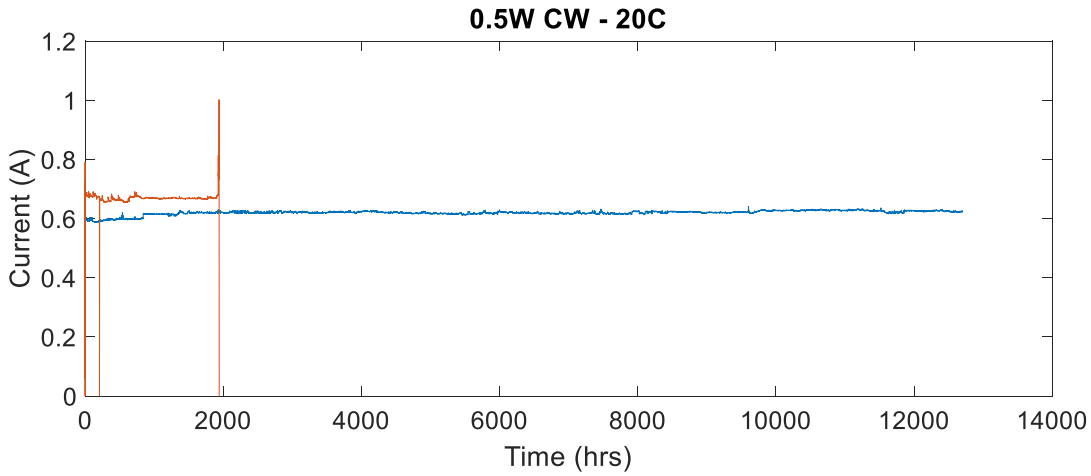


Fig 3.23 Shown in red and blue are devices operating at 0.5 W CW under constant power at a 20 °C submount temperature. The device in red failed at nearly 2000 hours and the device shown as the blue line has survived nearly 13,000 hours.

The device shown as the red line that failed ~2000 hours in the figure above failed because of outside conditions. This device appears to have failed from condensation. Shown in the figure below are optical images of the contact to the device.

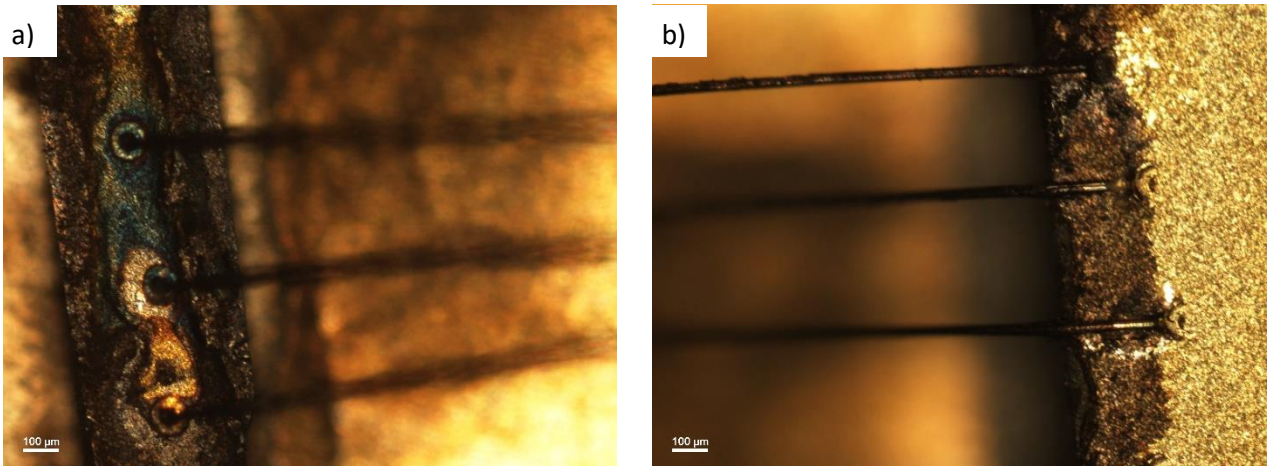


Fig. 3.24 (a) substrate side of a STA-RE QCL mounted epi-down on CuW with AuSn preforms operated under CW conditions with a constant power of 0.5W and submount temperature of 20 °C showing what appears to be scorch marks, and **(b)** Gold standoff showing similar scorch marks up the wirebonds and to the contact pad.

The optical images appear to show scorch marks bridging the wirebonds from the gold stand off to the substrate side of the device which would explain the jump in current prior to failure. To further run these devices at high output power with the large thermal dissipation requirements, a more controlled environment will be needed. This can be seen as either an environment where the humidity is monitored and controlled, or in a low-pressure environment. With all of the external failure mechanisms now mitigated, the ability to observe failures intrinsic to the device can be observed and characterized. With this comes the ability to create full-fledged lifetest studies which is an area for further study.

References

[1] J. Souto, J. L. Pura, M. Rodigues, J. Anaya, A. Torres, and J. Jimenez, “Mechanisms driving the catastrophic optical damage in high power laser diodes”, Proc. SPIE, Vol. 9348, 93480O, March 2015

[2] M. Ziegler, V. Talalaev, J. W. Tomm, T. Elsaesser, P. Ressel, B. Sumpf, and G. Erbert, “Surface recombination and facet heating in high-power diode lasers” *Appl. Phys. Lett*. **92**, 203506 (2008)

[3] F. Xie, H. Nguyen, H. Leblanc, L. Hughes, J. Wang, J. Wen, D. J. Miller, and K. Lascola, “Long term reliability study and life time model of quantum cascade lasers”. *Appl. Phys. Lett*. **109**, 121111 (2016)

[4] D. Botez, J. D. Kirch, C. Boyle, K. M. Oresick, C. Sigler, H. Kim, B. B. Knipfer, J. H. Ryu, D. Lindberg, T. Earles, L. J. Mawst, and Y. V. Flores. “High-efficiency, high-power mid-infrared quantum cascade lasers [Invited], *Opt. Mater. Express*. Vol. 8, No. 5, 1378, (2018)

[5] Q. Zhang, F. Liu, W. Zhang, Q. Lu, L. Wang, L. Li, and Z. Wang, “Thermal induced facet destructive feature of quantum cascade lasers”. *Appl. Phys. Lett*. **96**, 141117 (2010)

[6] V. Spagnolo, G. Scamarcio, D. Marano, H. Page, and C. Sirtori, “Thermoelastic stress in GaAs/AlGaAs quantum cascade lasers”, *Appl. Phys. Lett*. Vol. 82, No. 26, 4639-4641 (2003)

[7] Y. Qiao, D. Zhao, Y. Chen, J. Shao, H. Zhang, X. Tang, Y. Yuan, J. Li, Y. Zhao, Q. Ma, C. Guo, Y. Zhang, “Investigation of Electromigration Degradation Mechanism in Ti/Pt/Au Ohmic Contacts to p-GaAs Under High Current Density Stress”, *IEEE Trans. Electron Dev.*, Vol. 64, No. 11, 4581-4586, Nov. 2017

[8] Y. Sin, Z. Lingley, M. Brodie, B. Knipfer, C. Sigler, C. Boyle, J. Kirch, K. Oresick, H. Kim, D. Botez, L. J. Mawst, D. Lindberg, and T. Earles, “Degradation Mechanisms in MOVPE-Grown High-Power Buried-Heterostructure Quantum Cascade Lasers”. IEEE/OSA CLEO 2020.

[9] X. Liu, R. W. Davis, L. C. Hughes, M. H. Rasmussen, R. Bhat, and C. Zah, "A study on the reliability of indium solder die bonding of high power semiconductor lasers". *J. Appl. Phys*. **100**, 013104 (2006)

[10] A. Dirisu, G. Silva, Z. Liu, C. F. Gmachl, F. J. Towner, J. Bruno, and D. L. Sivco, "Reduction of Facet Reflectivity of Quantum-Cascade Lasers with Subwavelength Gratings". *IEEE Photon. Technol. Lett*., Vol. 19, No. 4, 221-223, Feb. 15, 2007.

CHAPTER FOUR

THERMAL & OPTICAL MODELING

4.1 Introduction

MOCVD is typically used in industry for QCL growths as the throughput is significantly higher when compared to MBE; however, it is still a time intensive and expensive process. To minimize unnecessary growths and optimize specific properties, both optically and thermally, it is critical that accurate models are implemented. As has been shown in Chapter 3, the thermal properties and the optimization of heat dissipation is of the utmost importance for reliable and efficient QCLs.

The following sections will be exploring thermal and optical modeling of QCLs using COMSOL Multiphysics.

4.2 Verification of the Model

Using COMSOL Multiphysics, 2D and 3D models of devices mounted in various configurations on submounts and with water-cooling have been developed. The overarching goal of these models is to extract the active-region temperature during CW, QCW, and pulsed operations. However, there are many variables that are unknown when trying to model a full QCL structure including the bondline thermal resistance, interface thermal resistance between layers such as InP and SiN, the surface morphology of the device, and the ridge uniformity to name a few. Verification of the model took place through a collaboration with Dr. Farzaneh at UW-Stevens Point.

This verification took place by correlating the thermal modeling in COMSOL with thermoreflectance measurements. The thermoreflectance measurement specifics can be found in

Becher's paper [1]. The device tested was a 40-stage, strain-compensated STA-QCL emitting near 5.0 µm, with 100 nm InGaAs confinement layers on either side of the active, 2 µm of 2E16 InP lower cladding, 3 µm of 2E16 InP upper cladding, and 1 µm of 2E19 contact layer. The devices were 9.5 µm-wide buried heterostructures with a 3 mm-long cavity, a HR-coated back facet, and an uncoated front facet. CCD-based thermoreflectance was used to measure the temperature rise of the uncoated front facet under QCW operation as a function of drive. The device was mounted epi-down on a copper heatsink with indium solder.

The devices were modulated with an ITC4005QCL under QCW operation with a pulse width of 35 µs pulse width and a duty cycle of 3%. Using thermoreflectance, the temperature rise of the active region as a function of time was measured at 5 µs intervals. The resulting temperature profile is shown in Figure 4.1.

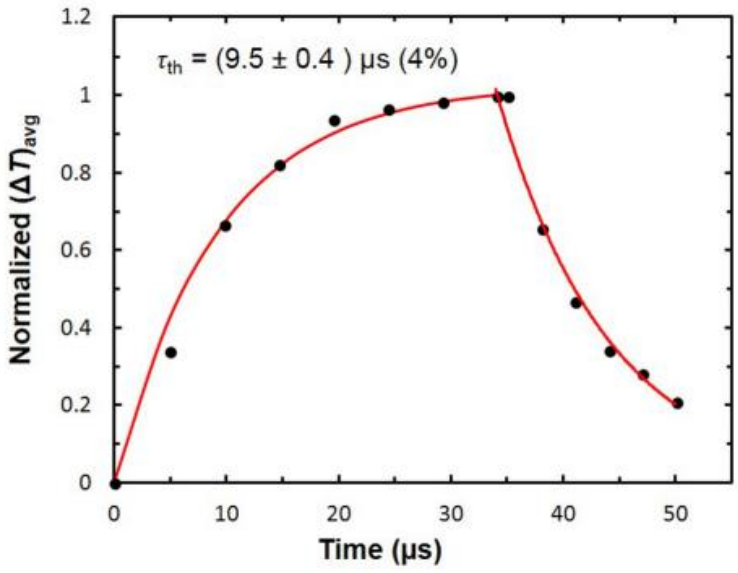


Fig. 4.1 Thermal response of a single-element BH QCL with a 35 µs pulse width at 1.1 A and 3% duty cycle. The black dots are the result of 20000 averaged frames from thermoreflectance at 5 µs intervals and the red line is the exponential fit. From the fit, the thermal time constant is found to be 9.5 ±0.4 µs.

From the data shown in Figure 4.1, the device reaches thermal saturation around 25 µs. Moving forward, at the device level, a 25 µs pulse width is used to essentially model the temperature rise in a device under CW operation. This allows for the device to experience the temperature rise under CW operation without the extreme heat-removal requirements that would otherwise be necessary.

With a suitable pulse width determined, the device heating as a function of drive was measured with thermoreflectance and in COMSOL. The pulse profile from the ITC4005QCL at 25 µs pulse width was measured and inputted into COMSOL. The drive as a function of time, coupled with the current-voltage characteristics for the specific device, allowed for the thermal transient response of the single element device to be modeled. The resulting thermal images of the active region as a function of drive from COMSOL and thermoreflectance are shown in Figure 4.2.

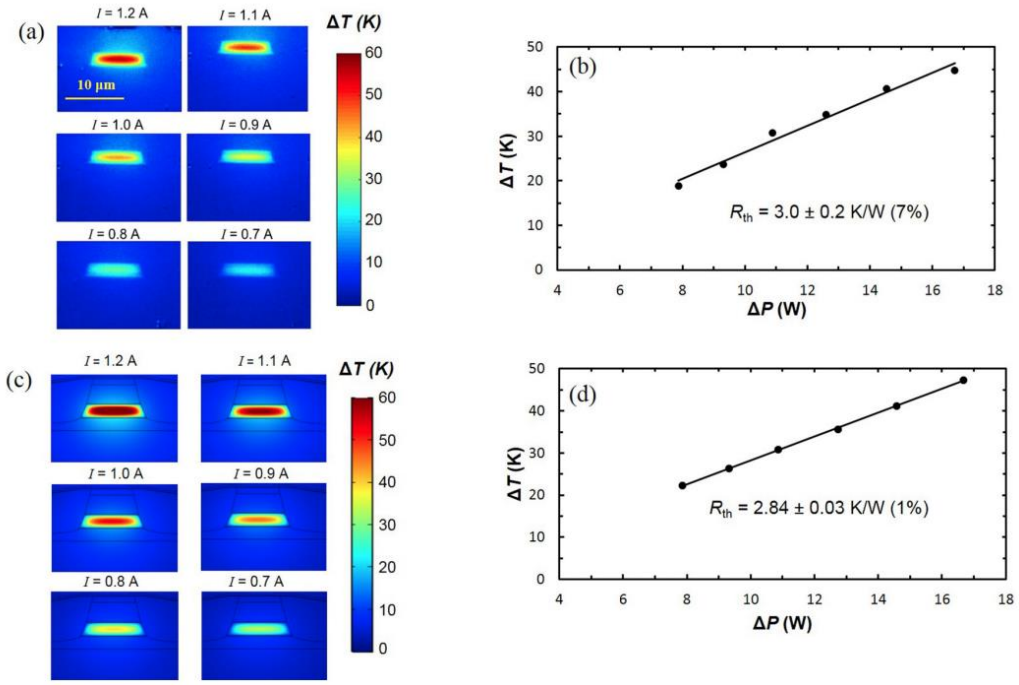


Fig. 4.2 Thermal images of the active region as a function of input drive for a 25 µs pulse at 3% duty cycle. **(a)** thermoreflectance measurements as a function of drive, **(b)** resulting thermal resistance extracted from thermoreflectance measurements, **(c)** COMSOL calculations as a function of drive, and **(d)** the

extracted thermal resistance calculation from the COMSOL calculations. ΔT is the temperature change averaged over the entire active region and ΔP is the change in input power minus the change in optical output power.

There is strong correlation between the simulated and experimental values for the both the thermal surface images as well as the extracted thermal resistance values. From Figure 4.2 (a) and (c) it is clear the heating within the active region is not uniform. This nonuniformity in heating leads to a thermal lensing effect. The change in temperature and refractive index both laterally and transversely, as a function of position and drive can be seen in Figure 4.3.

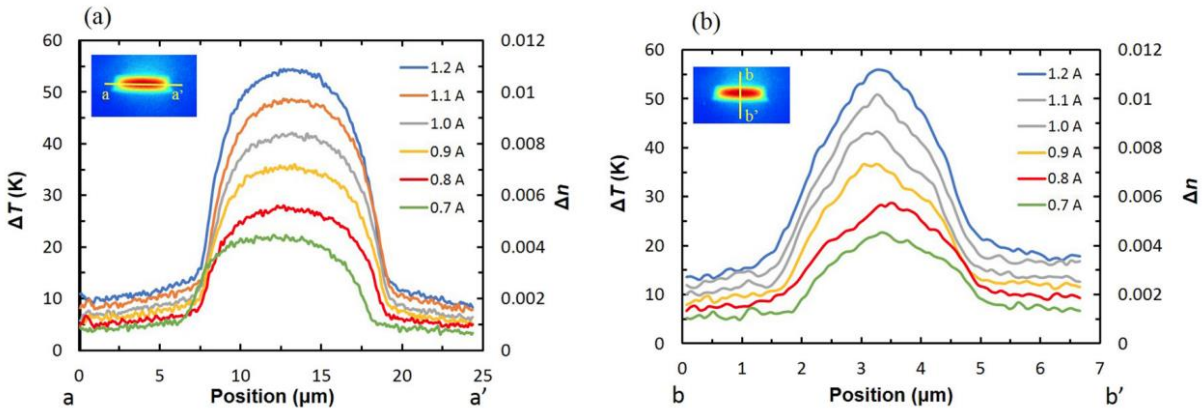


Fig. 4.3 Temperature profile as a function of position and drive. Insets show the direction of the scan. Each point within the scan is averaged over 5 pixels. **(a)** lateral temperature profile of a single element BH showing the temperature change on the left y-axis and resulting change in refractive index on the right y-axis and **(b)** the same output as (a), however, the direction of the scan is now transverse.

The change in refractive index relative is calculated using the following equation.

$$\Delta n = \frac{dn}{dT}\Delta T \tag{4.1}$$

where dn/dT is 2×10^{-4} K^{-1} [2]. This refractive index dependency on temperature has also been included in the COMSOL models such that thermal and optical modeling are interdependent.

With strong correlation between experimental and modeled values, the model has been verified to the point at which it can be used to show both thermal trends as well as the actual temperature rises. Current collaboration is in progress to further understand the thermal properties of these devices [3] and other material properties to further refine these models.

4.3 Optimized Geometries

Device temperature plays a large role in slope, threshold, maximum power, wall-plug efficiency, and especially the failure mechanisms of QCLs. While efforts to push to higher T_0 and T_1 values are currently ongoing through carrier-leakage suppression and efficient extraction designs, the active-region temperature rise remains a critical device parameter. The temperature rise in the active relative to the heat-sink can be well over 70 °C. Lowered active-region temperature rise will result in better performing devices and, ideally, fewer catastrophic failures.

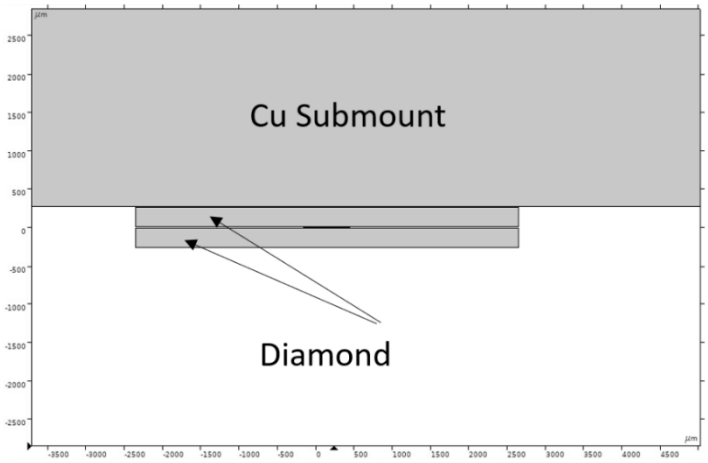


Fig. 4.4 A zoomed out image of a copper submount, with a diamond soldered to the surface, with a device indium bonded onto the diamond, followed by another diamond indium bonded to the substrate side of the device.

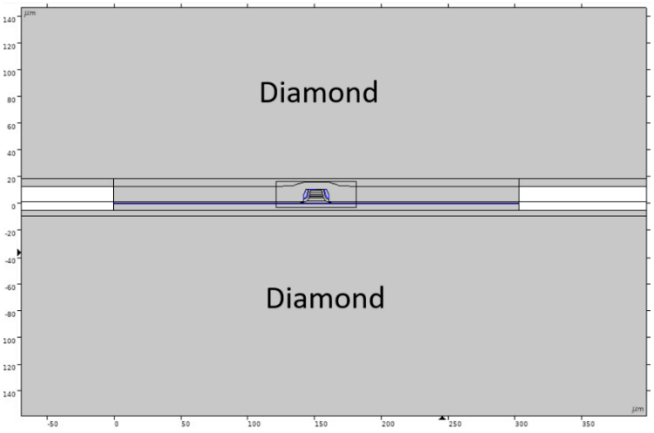


Fig. 4.5 A magnified image showing the device between two diamonds.

A current limitation of typical packaging techniques is that the heat removal can only occur via conduction to the heatsink while the radiation into the air on the surrounding 3 sides is extremely inefficient. Increasing the conduction surfaces by introducing a new highly thermally conductive material on the substrate side can significantly increase heat spreading and removal. This is schematically shown in Figure 4.5.

Through the exploitation of the extremely high thermal conductivity of diamonds ideally the active-region temperature rise will reduce significantly. To quantify the potential decrease in active region temperature, these structures have been thermally simulated in COMSOL. Shown in Table 4.1 are the input parameters for the model.

Table 4.1 COMSOL Multiphysics device parameters input where WPE is the wall-plug efficiency.

Input Power	20 W
WPE	10%
Cavity Length	3 mm
Ridge width	10 µm
Heatsink Temp.	20 °C

Holding these values constant, other parameters can be varied to develop a design with the lowest action-region temperature rise including, the spacer thickness, the Fe-InP regrowth thickness, as well as the waveguide design. To show how these designs would affect the heating, a baseline model must first be calculated. Figure 4.6 below shows the baseline design with 5 µm of Au-plating and indium bonding the substrate to a diamond.

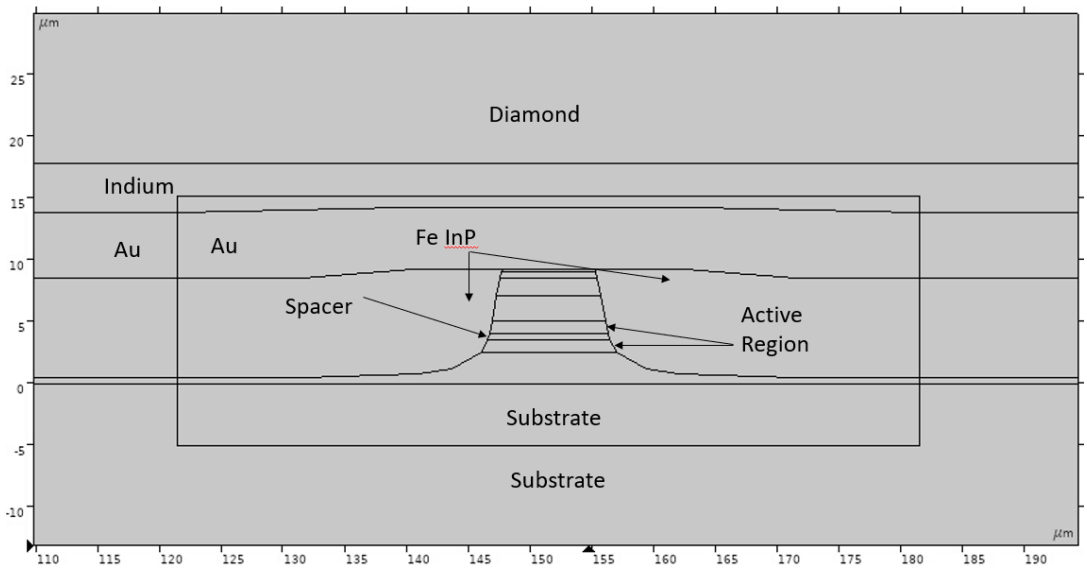


Fig. 4.6: QCL BH with 5 µm of Au plated on the epi-side and mounted epi-down on diamond with indium. Diamond is also bonded to the substrate side.

The average active region temperature rise is 67 °C. One important note, while the image shows a spacer of some thickness, for this 67 °C value, the spacer thickness was set to 0 µm. When varying the spacer thickness, this occurs at the location at which the spacer will be inserted, that is in the middle of the active region. With the baseline calculations, variations begin with varying the spacer and regrowth thickness.

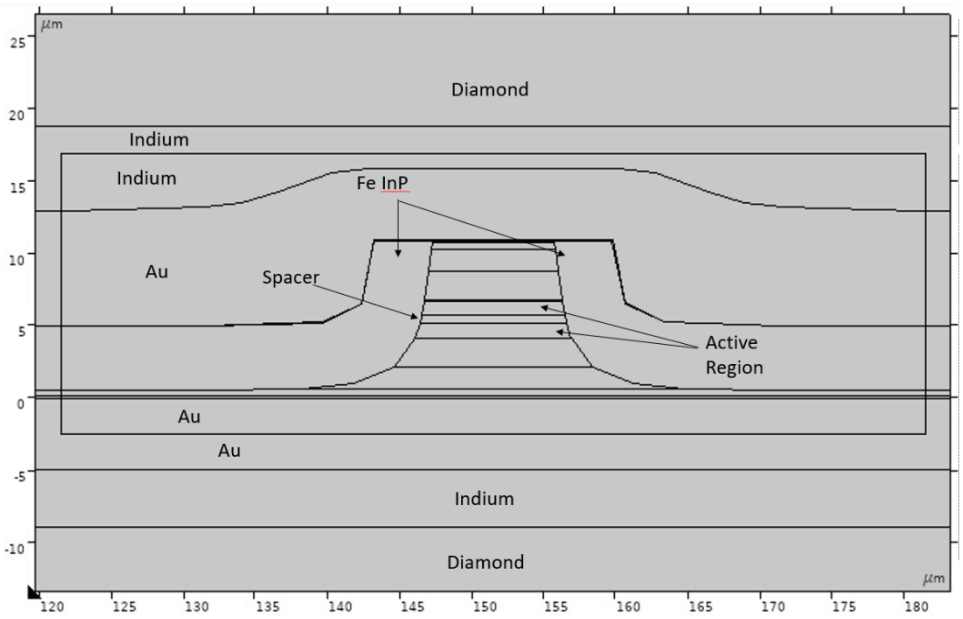


Fig. 4.7: BH with 5 µm Au-plating, epi-down with indium on diamond, with the substrate removed and bonded with indium to diamond. The regrowth thickness is 5 µm.

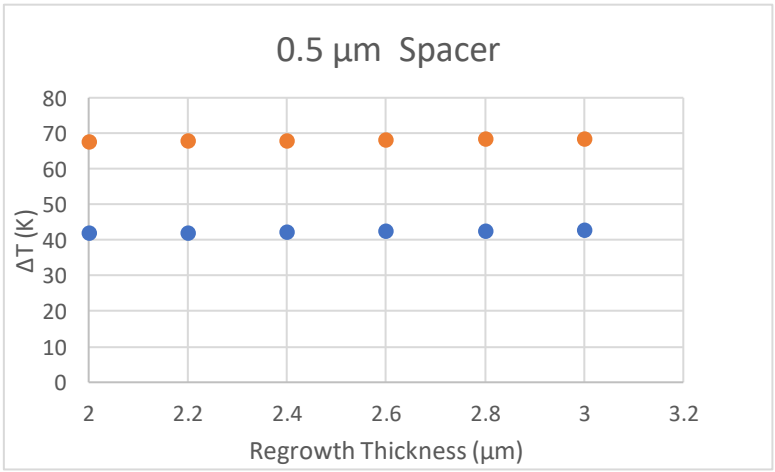


Fig. 4.8: Varying spacer thickness with a 5 µm-thick regrowth. The orange dots are the maximum temperature within the active region and the blue dots are the average values in the active region.

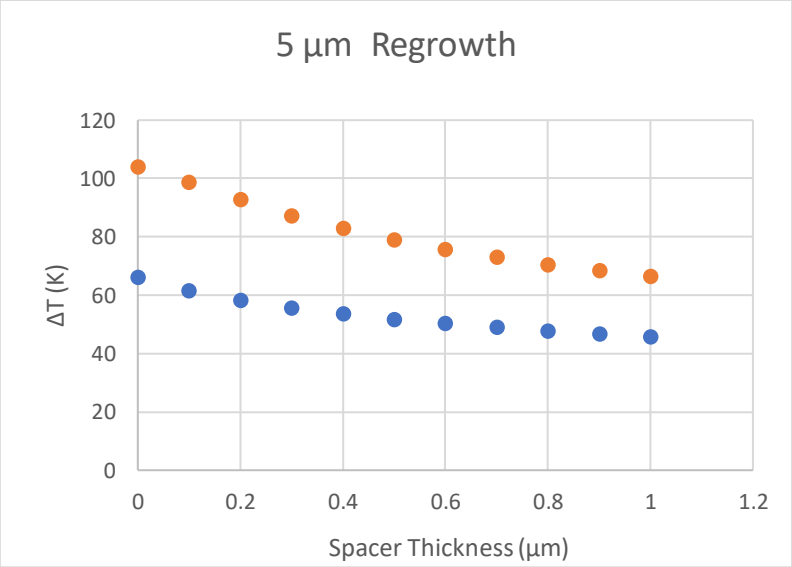


Fig. 4.9: Varying regrowth thickness while holding the spacer thickness constant at 0.5 µm. The orange dots are the maximum temperature within the active region and the blue dots are the average values in the active region.

Figures 4.8 and 4.9 show how the maximum temperature rise (orange) and the average temperature rise (blue) depend on the spacer thickness and regrowth thickness. The active-region temperature rise has a larger dependence on the spacer thickness than the regrowth thickness. By removing the substrate, introducing a 0.5 µm spacer, and decreasing the regrowth thickness from 10 µm to 2 µm, a decrease of 37% average core temperature rise is calculated. This 37% can be further increased by moving towards copper or silver plating instead of gold, further increasing the spacer thickness, or decreasing the regrowth thickness.

However, optical losses may be introduced by reducing the regrowth thickness to such extreme values. Shown below in Figure 4.10 is a design that may be used to keep the regrowth thin and maintain lower temperatures, while keeping the lossy gold further from the active region and the propagating mode.

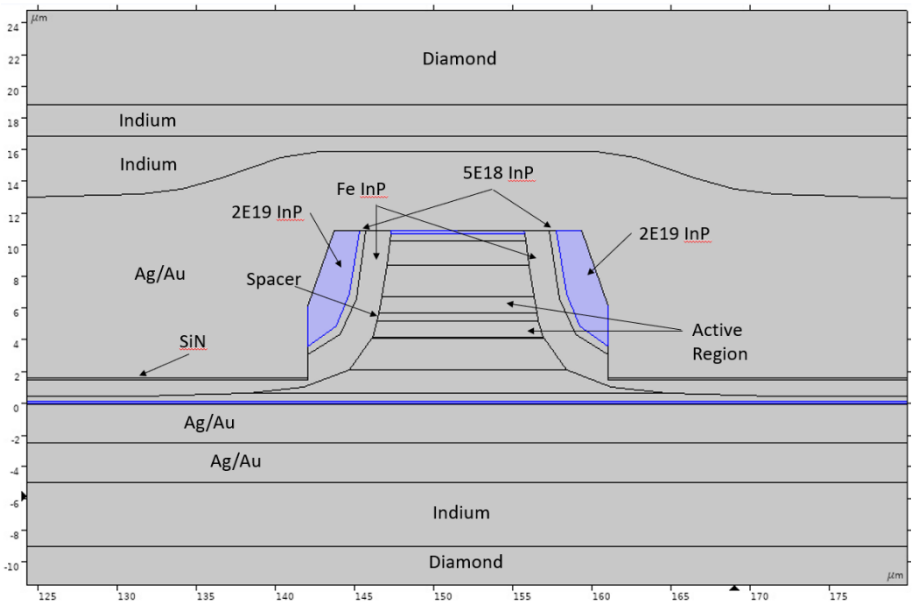


Fig. 4.10: BH variation to reduce optical loss while keeping the temperature low.

In the design shown in Figure 4.10, the regrowth is deeply etched and then followed by a thin SiN layer to act as the blocking layer in the field. To minimize optical losses to the plated metal layer, the Fe-InP, 5E18 InP, and 2E19 InP around the active remains. To keep the temperature low, the Fe-InP regrowth thickness at 2 µm as in the previous results is held constant and the spacer thickness is varied in this new configuration.

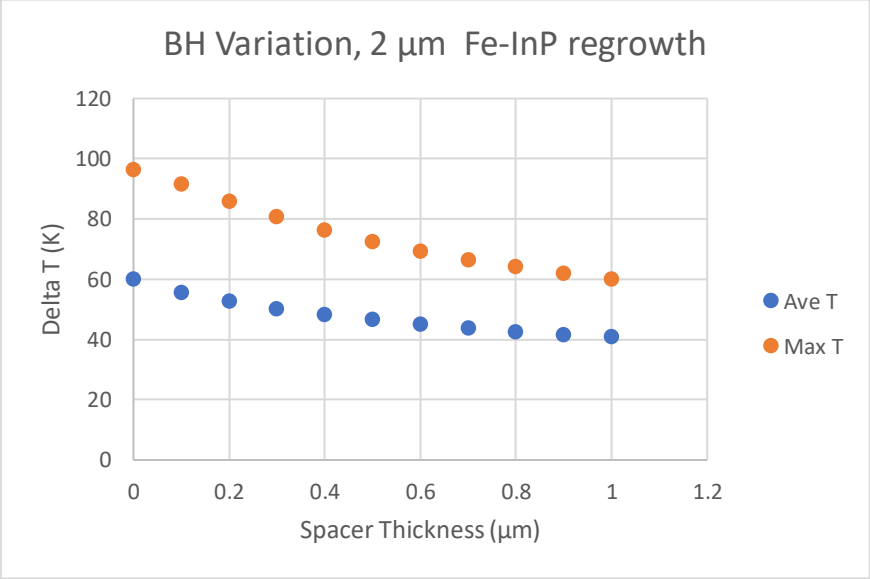


Fig. 4.11: 2 µm regrowth thickness while the spacer thickness is varied and while the BH is in the configuration shown in Figure 4.10. The orange dots are the maximum temperature within the active region and the blue dots are the average values in the active region.

This variation, results shown in Figure 4.11, show an average core temperature rise decrease of 31% when compared to the baseline. Although the temperature is not as low as the previously discussed design, it may prove to lead to lower optical losses. Further study on the optical losses of these designs is required.

4.4 T_0T_1 Empirical Model

A semi-empirical model was developed that is similar in nature to that developed by Suttinger [4] and expanded upon by Sigler [5]. Given a device structure and the accompanying pulsed and continuous-wave (CW) power-current-voltage (LIV) curves, the model can be populated with growth-specific values and the waveguide and device characteristics can be varied to determine the optimal design. Using COMSOL Multiphysics a 2D buried heterostructure (BH) epi-down on a diamond heatsink with indium solder is modeled to calculate the thermal dissipation, active-region temperature rise,

optical confinement factor, and effective index for varying device and waveguide geometries assuming the TM_{00} mode. The following equation is used to calculate the slope efficiency:

$$\eta_s = F_1 \frac{\hbar\omega}{q} \eta_i^d \frac{\alpha_m}{\alpha_m + \alpha_i} N_p \quad (4.1)$$

where η_s is the slope efficiency, η_i^d is the internal differential efficiency, α_m is the mirror loss, α_i is the internal/waveguide loss, F_1 is the fraction of mirror loss that is emitted from the front facet, and N_p is the number of stages. In the above equation, η_i^d and α_i have temperature dependencies and have a large impact on the slope efficiency as well as the maximum CW power which is why calculating the active-region temperature and self-heating is a critical step. This COMSOL simulation was run through a range of different efficiencies, stage thicknesses, ridge widths, number of stages, and bondline thermal resistances to generate a look-up table (LUT) that can be accessed via a Matlab script for different active-region designs. This COMSOL LUT contains all the thermal calculations for the average core temperature for various device geometries and with the characteristic temperature coefficients T_0 and T_1, how these temperature dependencies affect the device performance for this specific design can be calculated. Work is currently being performed on a model that predicts the T_0 and T_1 values.

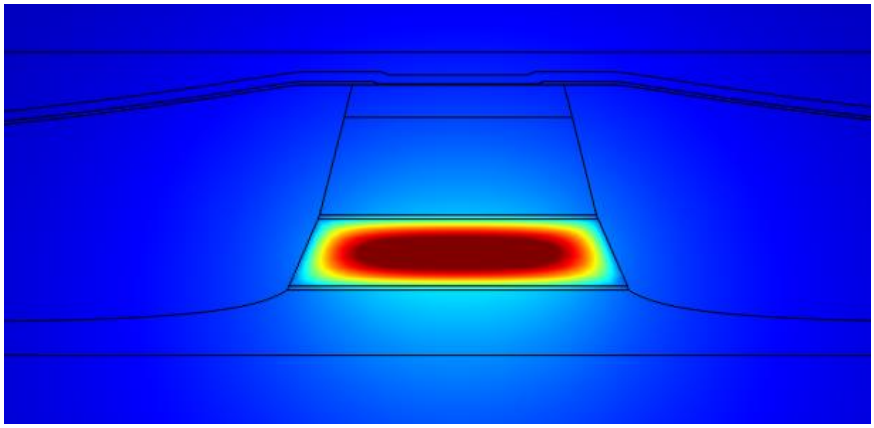


Fig. 4.12: BH thermally modeled in COMSOL epi-down on a copper submount mounted with indium with an 8 μm ridge-width.

If an automated process is desired, the output from nextnano including the expected current-voltage curve and the current density at rollover must be automated as well. With these values the experimental pulsed results would not be required to fit the semi-empirical model to and the ability to model the expected peak power, efficiency, vary the configuration, without the need to grow, fabricate, and test the structure would be available. With the implementation of this process the ability to vary the injector doping, device structure, and reflectivities is included to see the effect on device performance.

To use this method, however, T_0 and T_1 values are required for all potential designs. Currently there is no model that can predict the T_0 and T_1 values for quantum cascade lasers emitting in the mid-IR. We are working towards a model that includes the general temperature characteristics of the internal and backfilling losses and calculates the longitudinal optical (LO) phonon lifetimes, interface roughness (IFR) lifetimes, and the alloy disordering (AD) lifetimes, as a function of temperature. Therefore, an empirical model has been developed that incorporates the device temperature dependencies and fit those to previous device results for T_0 and T_1 values. The LO lifetimes and their temperature dependency is calculated using the following equation [6]:

$$\tau_i(T) = \frac{\tau_o}{1+\frac{2}{\exp\left(\frac{E_{LO}}{k_B T_e}\right)-1}} \tag{4.2}$$

Where τ_o is zero-temperature LO-phonon lifetime and E_{LO} is the energy of the LO phonon. Between the temperatures of 20 °C and 60 °C the decreases in LO lifetimes are ~9%. It was found that reducing this value to ~5% over the same temperature range achieved much better agreement with experimental values.

The equation for the backfilling loss is based off Maulini's calculation of the thermally excited backfilling concentration [7,8]:

$$n_{therm} = n_s \exp\left(\frac{\Delta_{inj}}{2k_B T_e}\right)\left(\frac{\sinh\left(\frac{\Delta_{inj}}{2N_{inj}k_B T_e}\right)}{\sinh\left(\frac{(N_{inj}+1)\Delta_{inj}}{2N_{inj}k_B T_e}\right)}\right) \tag{4.3}$$

With the thermally excited carrier concentration the effective backfilling-current density [9] and the equivalent backfilling loss [8] can be calculated using the following equations.

$$J_{th,BF} = \frac{q}{\tau_{up}} n_{therm} \tag{4.5}$$

$$\alpha_{BF} = g_c N_p n_{therm} \tag{4.5}$$

where,

$$g_c = \frac{4\pi e^2 z_{ij}^2 \Gamma}{\varepsilon_o n_{eff} \lambda 2\gamma_{ij} L_p} \tag{4.6}$$

where n_s is the sheet density doping, Δ_{inj} is the difference in energy between the lower laser level and the ground state in the injector of the following stage, T_e is the electronic temperature, N_{inj} is the number of states in the extractor, Γ is the confinement factor, $2\gamma_{ij}$ is the linewidth broadening, z_{ij}, is the dipole element between the states i and j, and L_p is the length of a single stage. It can be seen from this equation and Figure 4.13 (a), that the backfilling loss has a strong dependency on the electronic temperature.

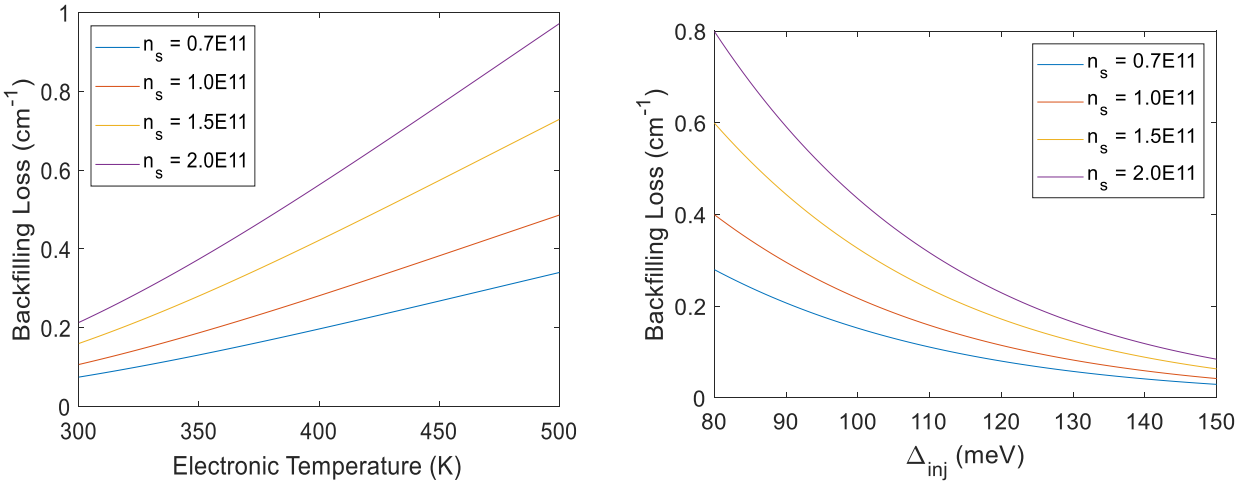


Fig. 4.13: (**a)** Backfilling loss dependency on electronic temperature with a constant Δ_{inj} of 120 meV. **(b)** Backfilling loss dependency on Δ_{inj} at a constant electronic temperature of 350 K.

With the backfilling loss handled, equation 4.7 shows how the empty cavity loss and intersubband loss is incorporated.

$$\alpha_w = \alpha_{empty\ cavity} + \alpha_{ISB} \tag{4.7}$$

where $\alpha_{empty\ cavity}$ is taken to be 0.5cm^{-1} and the α_{ISB} calculation is shown below:

$$\alpha_{ISB} = D_1 * \pi n_s 2\Upsilon_{ij} \exp\left(-\frac{E_{subband}}{5k_bT_{lattice}}\right) \tag{4.8}$$

$E_{subband}$ is the energy separation between the injector lower and upper minibands and D_1 is the empirical fitting factor. This empirical fitting factor is different at the temperatures of interest and calibrated to Design 5g_mod and then verified with a couple more designs. Design 5h will be shown in more detail. Through an iterative process the threshold-current density, losses, electronic temperature, and internal efficiency are calculated until they converge as they are all interdependent. Figure 4.14 (a) and (b) show how the leakage changes as a function of heatsink temperature for Design 5g_mod.

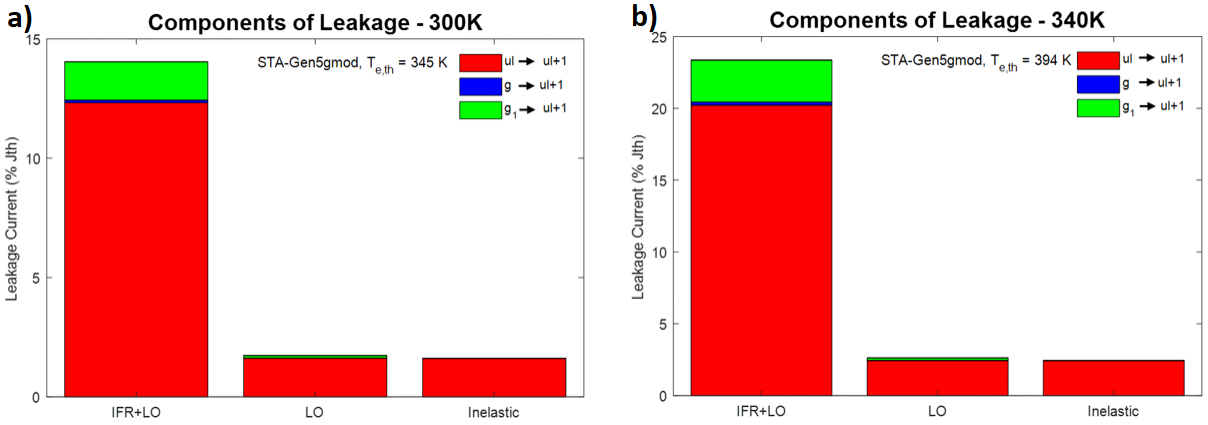


Fig. 4.14 (a) Leakage current for each state to the state above the upper lasing level at 300 K for 5g_mod and **(b)** the leakage current from each state to the state above the upper lasing level at 340 K for 5g_mod.

The leakage significantly increases at elevated temperatures which is a function of the losses, the leakage, and the electronic temperature. The two critical values we are extracting to estimate the T_0 and T_1 values are the threshold-current density and slope efficiency as functions of temperature. We use the following equation to extract the T_0 T_1 values.

$$J_{th}(T_{ref} + \Delta T) = J_{th}(T_{ref}) \exp\left(\frac{\Delta T}{T_0}\right) \tag{4.9}$$

$$\eta_s(T_{ref} + \Delta T) = \eta_s(T_{ref}) \exp\left(-\frac{\Delta T}{T_1}\right) \tag{4.10}$$

These results were used to calibrate the empirical fitting factor, D_1, to achieve T_0 and T_1 values of 226K and 287K which both have percent errors less than 10%.

Table 4.2: Compiled key values for Design 5g_mod with submount temperatures of 300 K and 340 K using the described T_0T_1 empirical model.

Submount Temp (K)	Electronic Temp (K)	J_{th} (kA/cm^2)	Slope (W/A)	α_w (cm^{-1})	α_{BF} (cm^{-1})	Leakage (% of J_{th})
300	344.8	1.38	3.55	2.8	0.175	14.04%
340	393.5	1.647	3.09	2.89	0.264	23.36%

To show that this can be used for other designs as well, this same method has been applied to a design that is quite different, Design 5h.

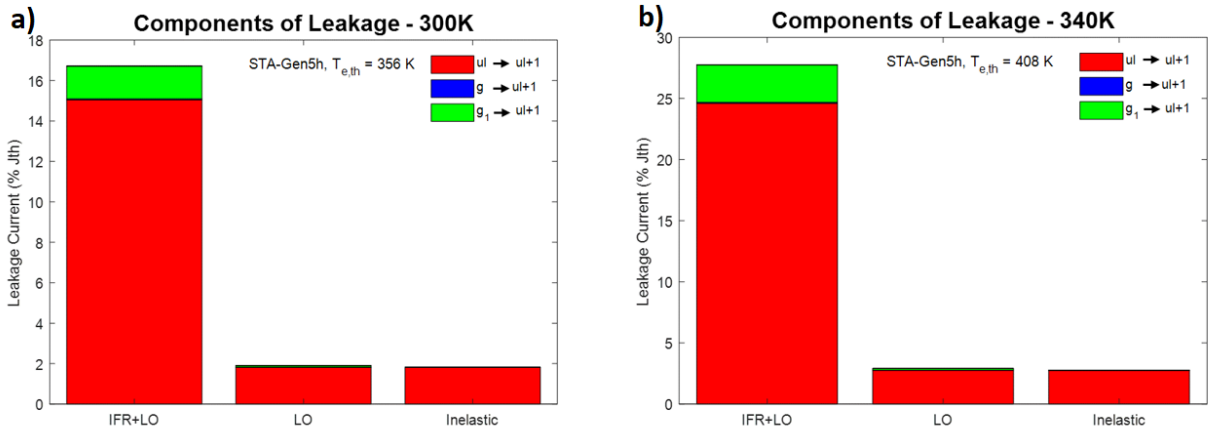


Fig. 4.15 (a) Components of leakage for Design 5h at 300 K, and **(b)** components of leakage for Design 5h at 340 K.

The extracted T_0 and T_1 values are 199K and 237K which again are within 10% error of the experimentally determined values.

Table 4.3: Compiled key values for Design 5h with submount temperatures of 300K and 340K using the described T_0T_1 empirical model.

Submount Temp (K)	Electronic Temp (K)	J_{th} (kA/cm^2)	Slope (W/A)	α_w (cm^{-1})	α_{BF} (cm^{-1})	Leakage (% of J_{th})
300	355.7	1.67	3.295	3.005	0.1228	16.71%
340	408.1	2.045	2.783	3.11	0.194	27.76

For both cases, the modeled threshold-current density is nearly identical to experimental values and the slope-efficiency calculation is consistently within 10% error. The slope calculation also verifies the calculated the waveguide and backfilling loss, and the distribution between the two.

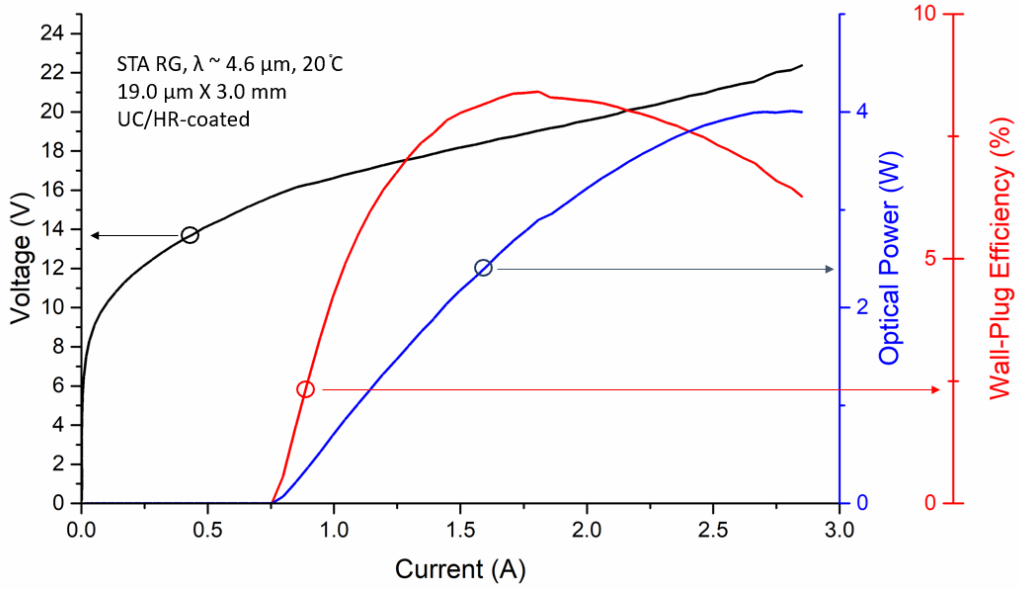


Fig. 4.16 Experimental results from a ridge guide fabricated with the base design of 5g_mod.

With the threshold -current density and expected slope efficiency, a CW LI projection can be calculated, and with nextnano, a full CW LIV can be obtained. An example using 5g_mod is presented, however, as the LIV is experimentally gathered, shown in 4.16, the use of nextnano here is not required. Figure 4.17 (a) and (b) show the CW power compares to the pulsed LIV for the ridge guide dimensions with an uncoated front facet and HR coated back facet.

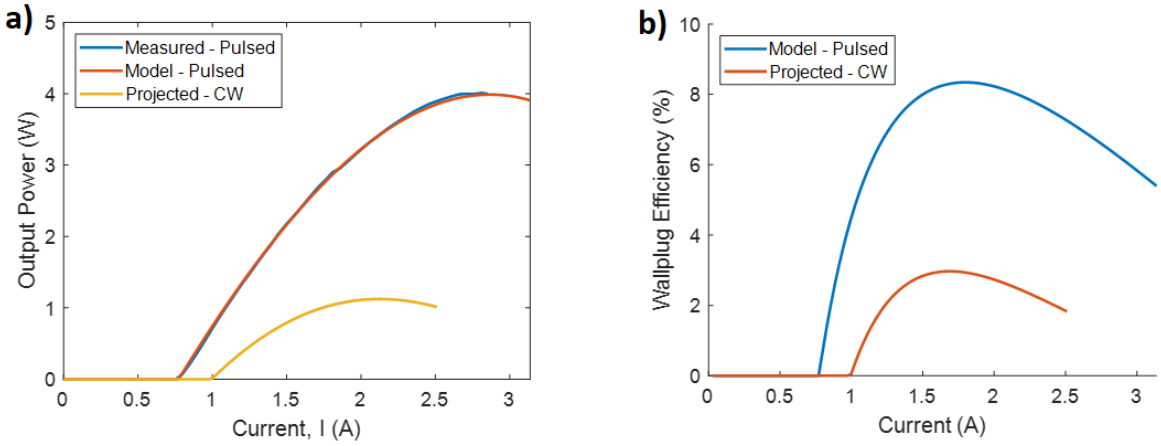


Fig. 4.17 (a) Using Figure 4.16 as a pulsed input, the CW performance is calculated and how it compares to pulsed operation, and **(b)** WPE of the projected CW power and how it compares to the same values under pulsed conditions.

As shown above, we can predict the CW performance of a device with the pulsed input. There is also the ability to vary the ridge width, length, number of stages, and coatings to determine the optimal configuration.

The above calculations are done under the assumption the device is a buried heterostructure with a 19 µm-wide ridge and a cavity length of 3 mm which is not ideal for heat dissipation. If the dimensions and coatings are changed, how this device with more optimal dimensions would perform under CW operation can be calculated. The following calculation assumes a ridge width of 6 µm, a cavity length of 7 mm, and LR/HR coatings with reflectivities of 0.10/0.95. Figure 4.18 (a) and (b) show how the pulsed and CW LIV's would change from pulsed to CW operation.

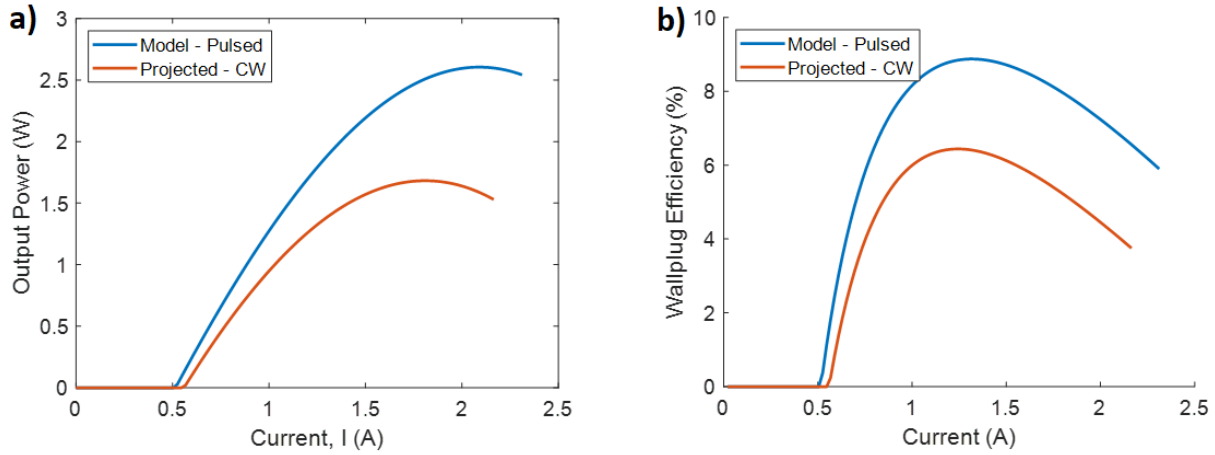


Fig. 4.18 (a) Calculated pulsed and CW LIVs assuming a 6 µm ridge, a 7 mm cavity, LR/HR coatings with reflectivities of 0.1/0.95, and the base material 5g_mod, and **(b)** Calculated pulsed and CW WPE assuming a 6 µm-wide ridge, a 7 mm-long cavity, LR/HR coatings with reflectivities of 0.10/0.95, and the base material 5g_mod.

As the device has a narrower ridge, there is better heat extraction which leads to higher output power. If we use the T_0T_1 values that were calculated using the above method as opposed to the experimental values, which was used in the above calculations, the peak power and wall-plug efficiency changes by less than 1%.

Moving forward for the device performance analysis, to make it entirely automated, the IFR/T_0T_1 calculations need to be incorporated into the semi-empirical CW script. All the necessary values are currently calculated in the scripts.

References

[1] N. Becher, M. Farzaneh, B. Knipfer, C. Sigler, J. Kirch, C. Boyle, D. Botez, L. J. Mawst, D. F. Lindberg III, and T. Earles, "Thermal imaging of buried heterostructure quantum cascade lasers (QCLs) and QCL arrays using CCD-based thermoreflectance microscopy", *J. of Appl. Phys*. **125**, 033102 (2019)

[2] C. Sigler, C. -C. Chang, J. D. Kirch, L. J. Mawst, D. Botez, and T. Earles, "Design of Resonant Leaky-Wave Coupled Phase-Locked Arrays in Mid-IR Quantum Cascade Lasers*", IEEE J. of Sel. Top. Quantum Electron*. **21**, 1200810 (2015)

[3] G. R. Jaffe, S. Mei, C. Boyle, J. Kirch, D. E. Savage, D. Botez, L. J. Mawst, I. Knezevic, M. G. Lagally, and M. A. Eriksson, "Measurements of the Thermal Resistivity of InAlAs, InGaAs and InAlAs/InGaAs Superlattices", *ACS Appl. Matls. & Interf*. 11(12) 11970-11975 (2019)

[4] M. Suttinger, R. Go, P. Figueiredo, A. Todi, H. Shu, J. Leshin, and A. Lyakh, "Power Scaling and Experimentally Fitted Model for Broad Area Quantum Cascade Lasers in Continuous Wave Operation," *Opt. Eng*., Vol. 57, No. 1, 011011, 2018.

[5] C. A. Sigler, "2-D Coherent Power Scaling of Mid-Infrared Quantum Cascade Lasers", Dissertation UW-Madison (2018)

[6] S. S. Howard, Z. Liu, D. Wasserman, A. J. Hoffman, T. S. Ko, and C. F. Gmachl, "High-Performance Quantum Cascade Lasers: Optimized Design Through Waveguide and Thermal Modeling" *IEEE Quantum Electronics*, Vol. 13, No. 5, 11054-11064, September 2007.

[7] R. Maulini, A. Lyakh, A. Tsekoun, and C. K. N. Patel, "$\lambda \sim 7.1$ µm quantum cascade lasers with 19% wall-plug efficiency at room temperature" *Optics Express* Vol. 19, No. 18, 17203-17211, August 2011

[8] Botez et al "Temperature Sensitivity of the Electro-Optical Characteristics for Mid – Infrared (λ = 3-16μm) – Emitting Quantum Cascade Lasers" *Journal of Physics D*: *Applied Physics* 49(4), 043001, December 2015.

[9] Yang et al. "Wall-plug efficiency of mid-infrared quantum cascade lasers". *J. Appl. Phys*. 111, 053111 (2012).

CHAPTER FIVE

EFFECTS OF CARRIER LEAKAGE & ITS SUPPRESSION VIA IFR-ENGINEERING

5.1 Introduction

While the first experimental CW QCLs was demonstrated only a year after the first fabricated device, peak wall-plug efficiencies of room temperature are still below 30% [1]. This is significantly lower than the 60% and upwards wall-plug efficiencies of diode lasers [2,3]. The increase of the QCL wall-plug efficiencies can be done through band-structure engineering to significantly reduce electron leakage and thermal backfilling (i.e., reduce the part of population of the lower state due to thermally excited carriers). The following sections will go through the various active-region designs that have been implemented previously, a novel method of designing active regions, and then the addition of a recently discovered mechanism to more accurately model various active-region designs.

5.2 Active-Region Design

QCL active-region design is an active area of research as the physics behind carrier scattering and leakage mechanism continues to be studied. Coupled with the temperature dependence, active-region design has progressed through a series of schemes to minimize leakage, promote efficient extraction, and maximize population inversion. Various active-region designs are discussed.

5.2.1 Conventional Structures

Conventional structures, as the name suggests, were based on the original design used for QCLs. That is, initially QCL structures had to be grown via MBE before MOCVD growth advancements allowed for QCL growth. While MBE can yield superb material quality, it lacks both industrial scalability and, most importantly, the ability to quickly change layer compositions. Conventional structures are characterized by wells and barriers of fixed compositions throughout the active region, as schematically shown in Figure 5.1.

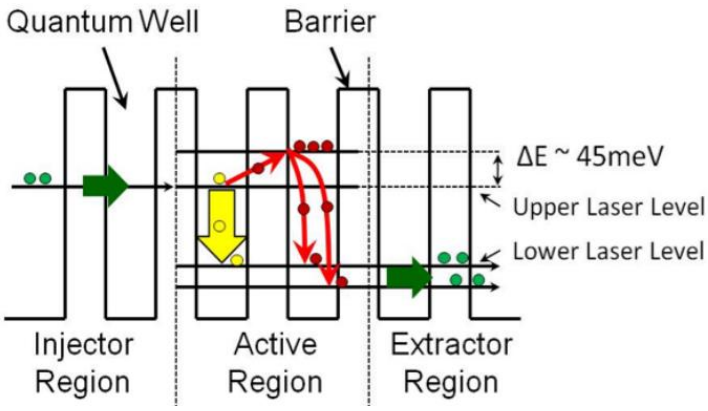


Fig. 5.1 Conduction-band edge for a conventional type QCL structure illustrating the various regions within a QCL core region, and a schematic representation of carrier leakage within and extraction from the active region. The barrier heights and well depths are constant, illustrating an alternating bi-layer of constant compositions of InGaAs and InAlAs.

Conventional structures are still in use for both research and industrial applications for their ease of growth and simulation. However, with the improvements in MOCVD growth, the ability to have both layer-thickness and composition agility has led to strong carrier-leakage suppression and high experimental internal efficiencies [1,4] while also achieving industrial scalability.

5.2.2 Nonresonant Extraction

Nonresonant extraction (NRE) works by no longer employing the two-phonon resonance condition that was previously employed for efficient carrier extraction at the lower laser level. This method employs a fifth quantum well in the active region which introduces a 1' state and pushes the energy difference between state 2 and state 1, $E_{21,}$ to a higher value. With the new 1' state, the $E_{31'}$ transition is now increased to ~ 140 meV [5] which allows for highly efficient extraction, and an increase in the energy difference between the upper laser (*ul*) level and the next-higher active-region state (i.e., the *ul +1* level) from 45 meV in conventional 4.5-5.0 μm-emitting QCLs [1] to ~ 60 meV. In addition, the lasing transitions are vertical which leads to low threshold-current density, J_{th}. However, the lower level (*ll*) lifetime is moderately high (~ 0.3 ps) which leads to lower than average (i.e. < 80 %) lasing-transition efficiency, and there is still significant carrier leakage due to strong overlap at interfaces between the *ul* and *ul +1* level wavefunctions moduli [6]. The carrier leakage is attested by relatively low T_0 and T_1 values (i.e., ~ 160 K and ~ 290 K) [7]

This design was shown to reach a CW WPE of 12.8 % for a single-facet emission at 4.6 μm [5] primarily because of low J_{th} values (~ 0.9 kA/cm^2), and a low waveguide loss (0.5 cm^{-1}) typical of MBE-grown devices.

5.2.3 Tapered Active – Deep Well

The NRE design was geared towards efficient extraction and low J_{th} values . The tapered active (TA) design focused on decreasing the carrier leakage. This is achieved by implementing a linear increase in the barrier height in the active region from the injector to the extractor region [8]. Increasing the barrier heights results in a significantly larger E_{54} energy, on the order of a 30% increase. State E_5 is the main carrier leakage path [9]. Decreasing this leakage also has a large effect on the J_{th} value. The band diagram for this structure is shown in Figure 5.2.

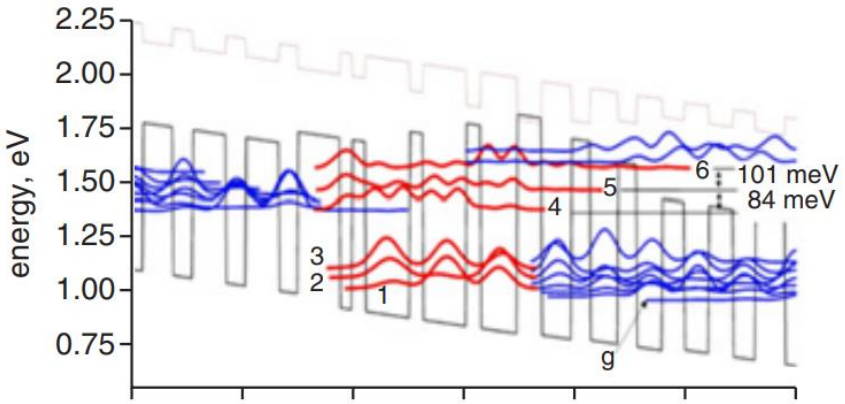


Fig. 5.2 Tapered Active – Deep Well (TA-DW) band structure showing a linear taper in the active region [8].

This device design resulted in a higher T_0 and a drastically higher T_1, 230K and 797K respectively. The experimental WPE was not reported but was projected to be > 20% from a single facet for the 4.6 µm to 4.8 µm range [9].

5.2.4 Step-Tapered Active

Like the tapered active scheme, the goal for the step-tapered active (STA) design is to increase the E_{54} separation and increase the lifetime τ_{54} in order to minimize carrier leakage. The STA structure increases the E_{54} separation, past what was achievable by the tapered-active (TA) structure, partly by introducing a significant degree of asymmetry [10]. With this design the E_{54} value reaches almost 100 meV leading to an almost negligible leakage current. The other reason the E_{54} separation increases is related to the Stark effect. Normally as the drive is increased the E_{54} value decreases due to the Stark effect [10]; however, by having a step-tapered design the Stark effect is minimized [10]. The band structure for a typical STA design is shown in Figure 5.3.

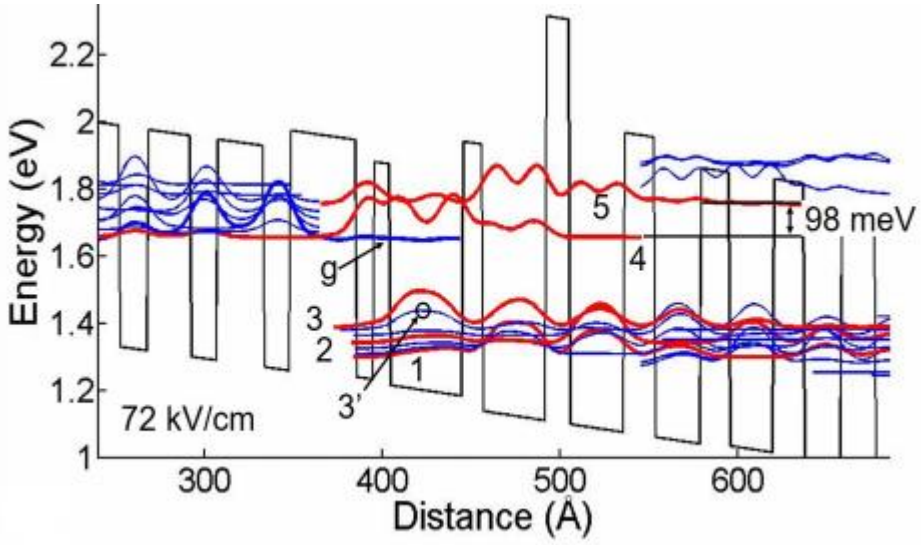


Fig. 5.3 Schematic representation of STA-QCL band structure and key wavefunctions

This structure has demonstrated, at ~5 µm emitting wavelength, high T_0 and T_1 values: 226 K and 653 K, respectively, for low-doped devices, as well as: 216 K and 400 K, for moderately-highly doped devices [4].

5.3 IFR-Engineered Design

As discussed in Chapter 2, IFR scattering plays a large role in QCL device physics. For reference, the IFR scattering rate is shown in the equation below.

$$\left.\frac{1}{\tau_{ij}^{IFR}}\right|_k = \frac{\pi m^*}{\hbar^3}\Delta^2\Lambda^2\delta U_k^2(z_k)\varphi_i{}^2(z_k)\varphi_j{}^2(z_k)e^{-\frac{\Lambda^2 m^* E_{ij}}{2\hbar^2}} \quad (5.1)$$

Important to note are the assumptions behind the above equation:

1. The roughness is fully correlated in the growth direction within each interface
2. Separation between layers is much larger than the interdiffusion length
3. Roughness of different interfaces are entirely uncorrelated

While QCLs do contain extremely thin layers, for the case of MOCVD, layers thinner than 1 nm are typically considered to be too thin to be achieved. From a design perspective, layer thicknesses are often targeted to be 1 nm or greater in thicknesses. Previous studies have looked at the interdiffusion length for MOCVD QCL material systems [11]. If the thinnest layer of 1 nm is assumed, and the interdiffusion length to be 0.5 nm, the separation between layers is larger than the interdiffusion length.

Looking at equation 5.1 from a design perspective, a few key parameters can be further analyzed to reduce the IFR scattering rate. The two most apparent parameters that can be reduced are the Δ and Λ parameters and while this would have a significant impact on IFR scattering rates, these are outside the scope of the design, and instead are currently being pursued through growth optimization. However, there are three parameters that can be optimized for the reduction of IFR scattering:

1. $\delta U_k^2(z_k)$ – conduction band offset
2. ${\varphi_i}^2(z_k){\varphi_j}^2(z_k)$ – overlap function at interfaces
3. $e^{-\frac{\Lambda^2 m^* E_{ij}}{2\hbar^2}}$ – exponential dependence on the energy separation

Using Figure 5.4 for reference, these 3 different parameters can be calculated and optimized for reduced scattering.

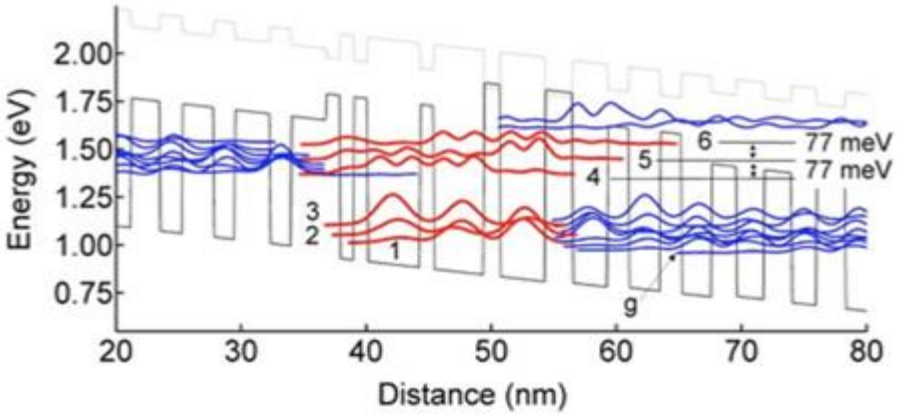


Fig 5.4 The conduction band edge for an TA design QCL emitting in the mid-IR. The various regions of a QCL active are labeled as well as the InAlAs barriers and InGaAs wells. State 5, the state above the upper laser level (*ul+1*), state 4, the upper laser level (*ul*), and state 3, the lower laser level (*ll*) are identified.

IFR scattering occurs at each interface. The scattering rate has a square dependence on the conduction band offset. Near 50 nm in Figure 5.4, there is a tall InAlAs barrier, based solely on the conduction band offset. It can be assumed that the IFR scattering is potentially high at this interface. Another key parameter is the overlap between the two wavefunctions of interest. The majority of IFR leakage is from *ul* to *ul+1* [6]. To reduce the leakage from *ul* to *ul+1*, a reduction in the sum of wavefunctions overlap at interfaces would definitely help. This can be viewed by pushing the *ul* level to the left while trying to push *ul+1* more towards the right, in reference to figure 5.4. The final method is increasing the energy separation between these two levels (*ul* and *ul+1*). From Figure 5.4 it can be seen the energy difference between *ul* and *ul+1* (states 4 and 5) is 77 meV. Figure 5.5 below shows how the IFR scattering rate varies with energy separation between the two states of interest when holding all else constant.

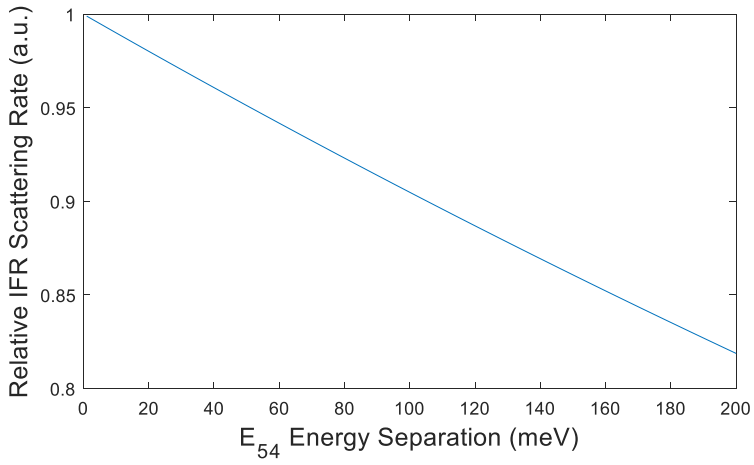


Fig. 5.5 Relative IFR scattering rate dependence on energy separation using the exponential dependence described in equation 5.1 assuming all else is held constant outside of the energy separation.

Considering these three parameters to minimize scattering rates, IFR-engineered designs can begin by minimizing the conduction band offsets, minimizing the overlap of the *ul* with *ul+1* energy levsls, and increasing the energy separation between *ul* and *ul+1*.

However, the optimization of QCLs is not as straightforward as simply minimizing the IFR scattering. Instead, the focus is on the maximization of the internal efficiency. The internal efficiency is defined below in equation 5.2.

$$\eta_i = \eta_{inj}\eta_{tr} \tag{5.2}$$

Where η_{tr} and η_{inj}, defined below, are the transition efficiency and injection efficiency, respectively.

$$\eta_{inj} = \left(1 - \frac{J_{leak}}{J_{th}}\right) \tag{5.3}$$

$$\eta_{tr} = \frac{\tau_{up,g}}{\tau_{up,g} + \tau_{ll,g}} \tag{5.4}$$

where J_{leak} is the leakage-current density and J_{th} is the threshold-current density. The injection efficiency is essentially the percentage of carriers that are not leaked and are available for the lasing transition (Note that the tunneling-injection efficiency η_{inj}^{tun} is taken to be unity, since generally η_{inj}^{tun} has values in the 0.96-0.98 range. $\tau_{up,g}$ is the effective upper-level global lifetime [10] and can be calculated with the following equation.

$$\tau_{up,g} = \tau_{4g}\left(1 - \frac{\tau_{3g}}{\tau_{43g}}\right) \tag{5.5}$$

where τ_{4g} and τ_{3g} are the global lifetimes for state 4 and state 3, respectively, τ_{43g} is the global lifetime for the 4 → 3 transition. In the past, QCL calculations have typically only considered LO-phonon scattering for all lifetimes and for carrier leakage [9] . However, now with the addition of IFR scattering in these global lifetimes and in the carrier-leakage process the ability utilize IFR engineering is presented and offers new insights into the leakage mechanisms. The total, final equation for the internal efficiency is shown below in equation 5.6.

$$\eta_i = \left(1 - \frac{J_{leak}}{J_{th}}\right) \frac{\tau_{up,g}}{\tau_{up,g} + \tau_{ll,g}} \tag{5.6}$$

It is clear from equation 5.6 that the route towards maximizing the internal efficiency is two-fold: increase carrier leakage suppression and minimize the $\tau_{ll,g}/\tau_{up,g}$ ratio. As discussed in the previous section, STA designs are optimal because, as one can see from the experimental T_1 values, it has strong leakage suppression by increasing the E_{54} energy spacing using the tall 3rd barrier in the active region. However, this tall barrier, and in turn large conduction band offset, has the potential for strong IFR scattering and leakage. Using the in-house developed scattering script, the leakage at each interface can be examined. Figure 5.6 shows how the leakage varies by interface within a QCL active region.

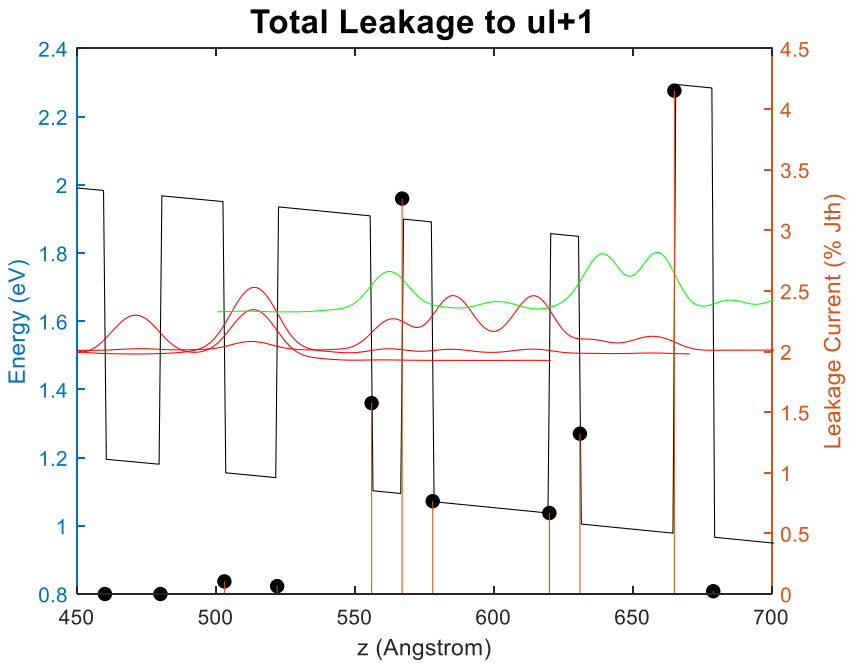


Fig. 5.6 Conduction band diagram for a STA-type QCL emitting in the mid-IR. The green wavefunction represents *ul+1*, or state 5, and the red wavefunctions from bottom to top are g_0, g_1, and *ul* or state 4, respectively. The left y-axis corresponds to the conduction band energies, the right y-axis corresponds to the stem graph illustrating the amount of leakage at each interface to *ul+1*, normalized with respect to the threshold-current density, and the x-axis is a function of position transverse to the growth direction.

It is clear from the figure above that not all interfaces contribute equally to leakage. The upstream side of the third barrier contributes over 4% of the threshold current density in leakage. The other main contribution comes from the first barrier in the active, which contributes nearly 3.5% of the threshold-current density in leakage. The first barrier has a significant contribution largely due to the strong overlap, at interfaces, with *ul*, while the third barrier has a significant contribution because it has a large conduction- band offset and has significant overlap. IFR engineering can be employed to mitigate these causes by either pushing the wavefunctions such that the overlap is not as significant and/or reducing the conduction band offset.

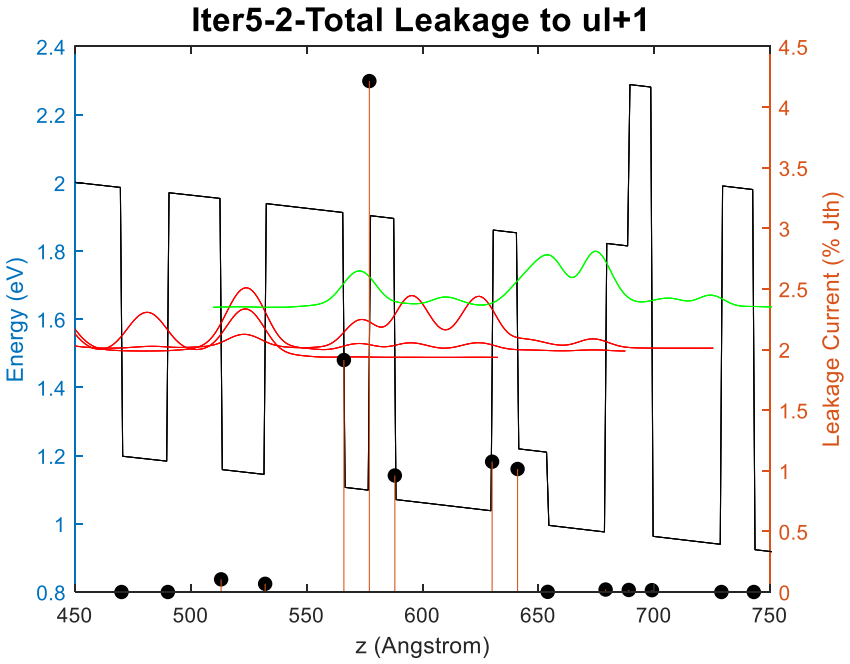


Fig. 5.7 Conduction band diagram for a QCL active region showing how the introduction of stepped barriers affects the IFR leakage.

As seen from Fig, 5.7, implementation of stepped barriers results in a 15% reduction is calculated IFR leakage which results in a 2% increase in internal efficiency. Through further refinement this is clearly a method that can be used to further increase device efficiencies and reduce leakage.

5.4 Graded Interfaces

Recent studies by Grange and Mukherjee [12,13] have introduced a new method of incorporating graded interfaces in the calculation of IFR through the incorporation of an axial correlation length, as discussed in Chapter 2. However, graded interfaces are now to be incorporated into the remaining parameters as well. Below is a list of parameters that are affected by graded interfaces:

1. Lattice constant
2. Conduction band edge
3. AD scattering (discussed in Chapter 2)

4. IFR scattering (discussed in Chapter 2)

The first parameter changed is the conduction band as it has the largest impact on the resulting wavefunctions. Shown in Figure 5.8 is an ungraded conduction band, and then Figure 5.9 shows the inclusion of a graded interface that follows an error function profile with an interfacial mixing width 'L' of 0.5 nm. The equation used for this interface is shown in equation 5.7.

$$c(z) = c_o + d_o \operatorname{erf}\left[\frac{2\sqrt{ln(2)}(z-z_o)}{L}\right] \tag{5.7}$$

where c_o and d_o are the offset and scale parameters, respectively, and z_o is the center of the interface.

The lattice parameter is graded in accordance with Vegard's law using equation 5.8 and using the quaternary, $Al_xIn_yGa_{1-x-y}As$

$$a = (y)a_{InAs} + (1-x-y)a_{GaAs} + (x)a_{AlAs} \tag{5.8}$$

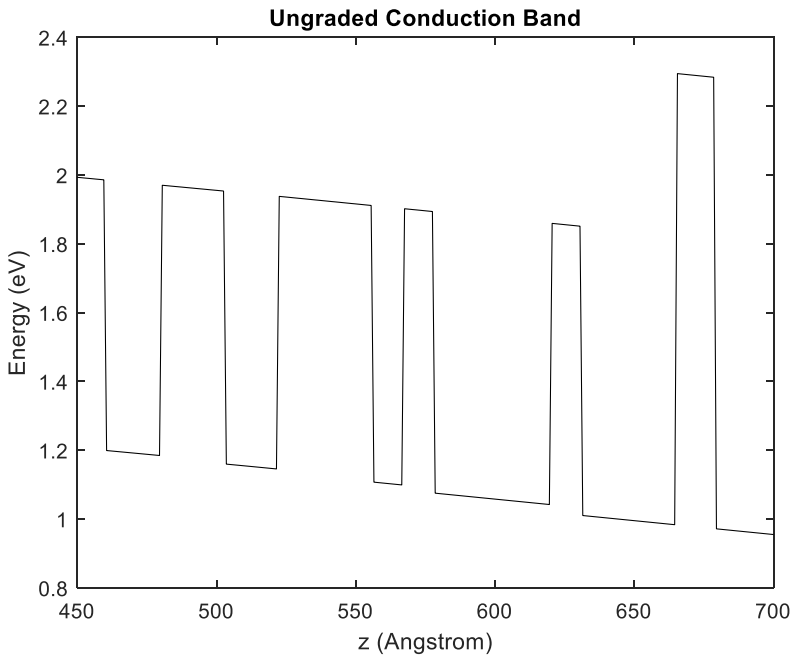


Fig. 5.8 Ungraded conduction band edge.

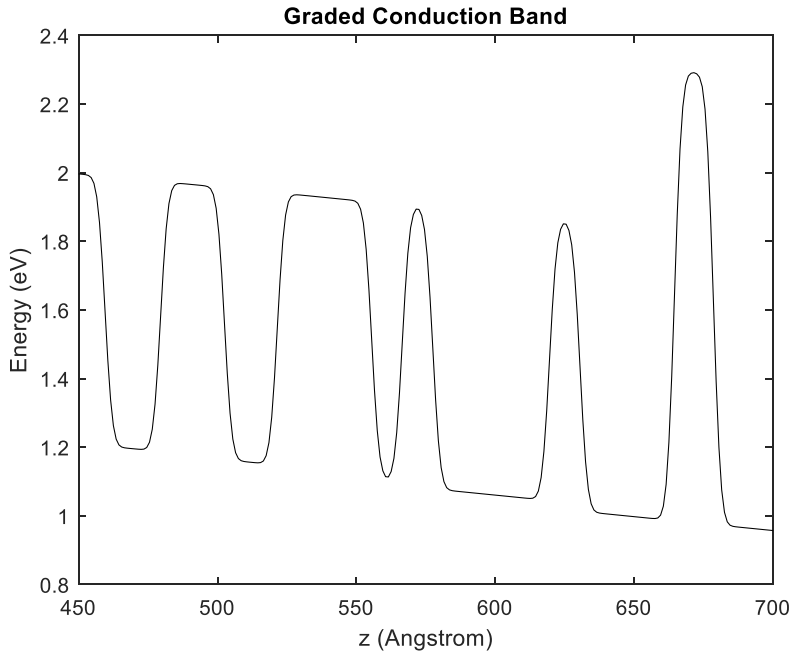


Fig. 5.9: Graded conduction band edge following the error function presented in equation 5.7, where the interfacial mixing width, L, is 0.5 nm.

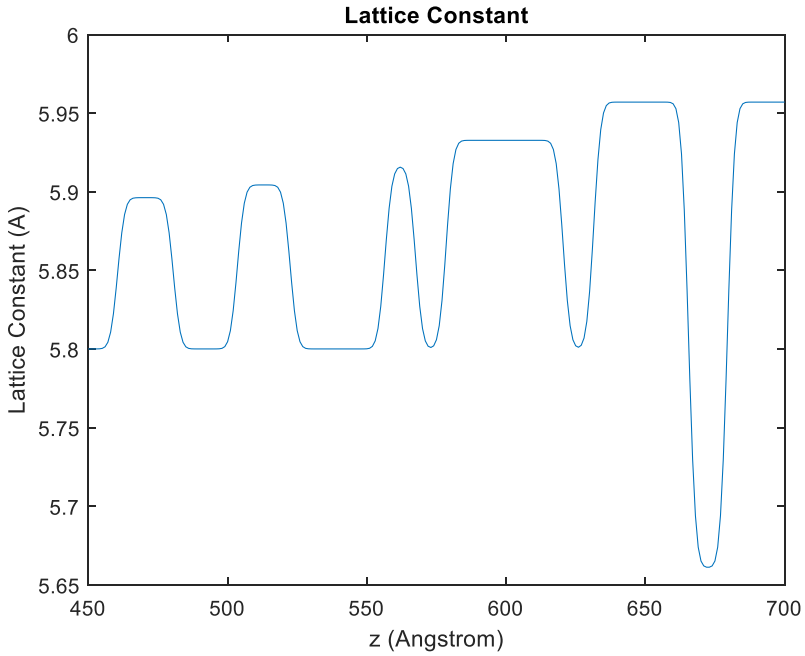


Fig. 5.10 Graded lattice constant through the active region based on equation 5.8.

Typically, the effective mass would also be graded, however, instead of using the effective mass as a function of position through the graded interface and through structure, the effective mass for each state is considered. For the AD- and IFR-scattering downward transitions, the effective mass used for the graded interface calculations is that for each individual state. In both cases, the effective mass considered for transitions between two states, is that for the state of lower energy. For IFR-scattering upwards transitions, that is, those involved in triggering carrier leakage from injector states and the *ul* level, both the initial state and final state effective masses are considered.

The addition of grading, the axial correlation length, and quaternary AD scattering plays a significant role in these scattering calculations. The following Chapter delves into where these values come from, how they were extracted from MOCVD-grown materials, and the effect they have on key device parameters.

References

[1] D. Botez, C.-C. Chang, L. J. Mawst, "Temperature sensitivity of the electro-optical characteristics for mid-infrared (λ = 3-16 µm)-emitting quantum cascade lasers," *J. Phys. D: Appl. Phys*. vol. 49, no. 4, 043001, 2016.

[2] M. Kanskar, D. Botez, T. Earles, T. Goodnough, L. J. Mawst, and E. Stiers, "73% CW Power Conversion Efficiency at 50 W from 970nm Diode Laser Bar", *Electron. Lett.*, 41 pp. 245-247, 2005.

[3]. Knigge, G. Erbert, J. Jonsson, R. Staske, B. Sumpf, M. Weyers, and G. Trankle, "Passively cooled 940 nm laser bars with 73 % wall-plug efficiency and high reliability at 98 W quasi-CW output power", Conference on Lasers and Electro-Optics, Europe, (2005)

[4] D. Botez, J. D. Kirch, C. Boyle, K. M. Oresick, C. Sigler, H. Kim, B. B. Knipfer, J. H. Ryu, D. Lindberg, T. Earles, L. J. Mawst, and Y. V. Flores. "High-efficiency, high-power mid-infrared quantum cascade lasers [Invited], *Opt. Mater. Express*. Vol. 8, No. 5, pp. 1378-1398, (2018)

[5] A. Lyakh, R. Maulini, A. Tsekoun, R. Go, C. Pflugl, L. Diehl, Q. J. Wang, F. Capasso, and C. K. N. Patel, "3 W continuous wave room temperature single-facet emission from quantum cascade lasers based on nonresonant extraction design approach," *Appl. Phys. Lett*., vol. 95, no. 14, 141113, 2009.

[6] C. Boyle, K. M. Oresick, J. D. Kirch, Y. V. Flores, L. J. Mawst, and D. Botez, "Carrier Leakage via Interface-Roughness Scattering Bridges Gap Between Theoretical and Experimental Internal Efficiencies of Quantum Cascade Lasers", Appl. Phys. Lett. **117**, 109901 (2020).

[7] A. Lyakh, R. Maulini, A. Tsekoun, R. Go, S. Von der Porten, C. Pfuegl, L. Diehl, F. Capasso and C. K. N. Patel, "High-performance continuous-wave room temperature 4.0-µm quantum cascade lasers with single-facet optical emission exceeding 2 W", *Proc. Natl. Acad. Sci. U.S.A*., vo1.107, no. 44, pp.18799–18802, 2010.

[8] J. Kirch, J. C. Shin, C. -C. Chang, L. J. Mawst, D. Botez, and T. Earles, “Tapered active-region quantum cascade lasers (λ=4.8 μm) for virtual suppression of carrier-leakage currents”, *Elec. Lett*. 48(4):234 (2012)

[9] D. Botez, S. Kumar, J. C. Shin, L. J. Mawst, I. Vurgaftman, and J. R. Meyer, “Temperature dependence of the key electro-optical characteristics for midinfrared emitting quantum cascade lasers” *Appl. Phys. Lett*. **97**, 071101 (2010)

[10] D. Botez, J. C. Shin, J. D. Kirch, C.-C. Chang, L. J. Mawst, T. Earles, “Multidimensional conduction- band engineering for maximizing the continuous-wave (CW) wallplug efficiencies of mid-infrared quantum cascade lasers”, *IEEE J. Sel. Top. Quantum Electron*., vol. 19, no. 4, 1200312, 2013; Correction: *IEEE J. Sel. Top. Quantum Electron*., vol. 19, no.4, 9700101, 2013.

[11] A. Rajeev, C. Weixin, J. D. Kirch, S. E. Babcock, T. F. Kuech, T. Earles, and L. J. Mawst, “Interfacial Mixing Analysis for Strained Layer Superlattices by Atom Probe Tomography”, *Crystals* **2018**, 8, 437

[12] T. Grange, S. Mukherjee, M. Montanari, L. Persichetti, L. Di Gaspare, S. Birner, A. Attiaoui, O. Moutanabbir, M. Virgilio, and M. De Seta, “Atomic-Scale Insights into Semiconductor Heterostructures: From Experimental Three-Dimensional Analysis of the Interface to a Generalized Theory of Interfacial Roughness Scattering”, *Phys*. *Rev. Appl.*, **13**, 044062 (2020)

[13] S. Mukherjee, A. Attiaoui, M. Bauer, O. Moutanabbir, “3D Atomic Mapping of Interfacial Roughness and Its Spatial Correlation Length in Sub-10 nm Superlattices”, *Appl. Mater. Interfaces* 2020, **12**, 1728-1739.

CHAPTER SIX

IFR-PARAMETER EXTRACTION AND IMPLEMENTATION

6.1 Introduction

Atom Probe Tomography (APT) is a quickly growing area of research for material characterization and analysis. It has played a large role in the investigation of defects, interstitials as well as many other material defects [1, 2]. It allows for the analysis of sub-nanometer structures in a full 3D reconstruction. This can be applied to the analysis of materials from precipitates and grain structures to interfaces. This is an important and critical tool for the analysis of interfaces as it allows for the investigation to be conducted in 3D across the entire interface as opposed to a 2D projection as would be seen with TEM. The main purpose of using APT here is to analyze the intermixing of elements at an interface and while with a sufficiently thin TEM sample the intermixing of atoms along an interface might be detected, extracting the elemental identity and spatial distribution of the atoms in 3D is substantially more difficult if possible at all.

APT evolved from field evaporative microscopy which is a more illustrative name. This technique is performed on extremely sharp needle-like samples with a typical tip radius of < 50 nm. The specimen is put close (~40 μm) to a counter electrode and exposed to an extremely high electric field (3-6 V/nm) where the applied voltage for the materials shown here range from ~1200V to ~4000V. While under this high electric field, the tip is pulsed with a laser of a specified frequency and energy. This additional energy imparted by the laser is enough to field evaporate atoms off the tip, ideally one-by-one. Then, based on time-of-flight mass spectroscopy and where the ion hits the detector, a 3D reconstruction of the atom positions and their chemical identity in the original tip can be created. To create accurate reconstructions,

a few other measurements are required for calibration; these can include SEM images of the tip before and after evaporation to determine the tip shape, high-resolution X-ray diffraction (HR-XRD) for layer thicknesses and compositions, as well as TEM measurements. In the following sections three groups of samples and APT runs with different reconstructions, calibrations, and analysis are investigated.

6.2 Samples Analyzed

There are three samples analyzed: 1) a thin superlattice (SL) structure with compositions that mimic some layers within relevant active-region designs, further referred to as the thin SL sample; 2) a full STA-QCL structure which has the same active-region design as described by Botez et al [3], further referred to as P17a; and 3) a full 40-stage strain-compensated 4.6 μm emitting STA-QCL structure, further referred to as 5g_mod. The purpose of studying the thin SL sample was to begin investigating thin layers of alternating InAlAs and InGaAs with compositions of relevance to state-of-the-art STA QCL active regions and how they evolve relative to the thickness of the layers.

The P17a structure, which has the same growth parameters as the SL samples, was analyzed to verify the thin SL results and to investigate whether there was thermally-induced interdiffusion at the layer interfaces by comparing the top and the bottom of the active region.

The purpose behind analyzing the 5g_mod sample was to analyze specific interfaces/layers that play key roles in the IFR/AD scattering and IFR-leakage calculations. The IFR parameters are experimentally extracted from isoconcentration surfaces through three interfaces of interest.

6.2.1 Thin SL Structure

A specific SL structure was grown on a (001)-oriented InP substrate by MOCVD in a close-coupled showerhead (3x2”) configuration at a temperature of 605°C, a reactor pressure of 100 torr, a 5 sec

interruption time between layers, and a susceptor rotation of 100 rpm. The V-III precursor ratios, group-III partial pressures, and growth rates are shown below in Table 6.1.

Table 6.1 V/III ratios and resulting growth rates used to grow the corresponding target composition.

Structure	*Target Composition*	*V/III Ratio*	*Group III Partial Pressure (Torr)*	*Growth Rate (nm/s)*
SL	$In_{0.44}Al_{0.56}As$	370	1.44E-3	0.11
SL	$In_{0.6}Ga_{0.4}As$	406	1.32E-3	0.10
Upper Cladding	$In_{0.52}Al_{0.48}As$	308	1.73E-3	0.13
Lower Cladding	$In_{0.53}Ga_{0.47}As$	476	1.12E-3	0.08

The sources used were: trimethylindium (TMIn), trimethylgallium (TMGa), trimethylaluminum (TMAl), phosphine, and arsine. The TMIn concentration was controlled in a feedback loop with an Epison concentration monitor. The targeted SL layers were 2x(1 nm:$In_{0.44}Al_{0.56}As$/1 nm:$In_{0.6}Ga_{0.4}As$); 2x(2 nm:$In_{0.44}Al_{0.56}As$/ 2 nm:$In_{0.6}Ga_{0.4}As$); 2x(3 nm:$In_{0.44}Al_{0.56}As$/ 3 nm:$In_{0.6}Ga_{0.4}As$) and 2x(4 nm:$In_{0.44}Al_{0.56}As$/ 4 nm:$In_{0.6}Ga_{0.4}As$) in the growth direction. The SL structure was bound on one side by 50 nm of $In_{0.53}Ga_{0.47}As$ towards the substrate and terminated by 50 nm of $In_{0.52}Al_{0.48}As$, followed by a thin InP-capping layer, shown schematically in Figure 6.1 (a). A TEM specimen was created via FIB and analyzed on-site, shown in Figure 6.1 (b). The measured layer thicknesses extracted from the TEM images were approximately 0.9 nm:$In_{0.44}Al_{0.56}As$/1.0 nm:$In_{0.6}Ga_{0.4}As$; 2.0 nm:$In_{0.44}Al_{0.56}As$/ 2.2 nm:$In_{0.6}Ga_{0.4}As$; 3.3 nm:$In_{0.44}Al_{0.56}As$/ 3.3 nm:$In_{0.6}Ga_{0.4}As$, and 4.3 nm:$In_{0.44}Al_{0.56}As$/ 4.1 nm:$In_{0.6}Ga_{0.4}As$.

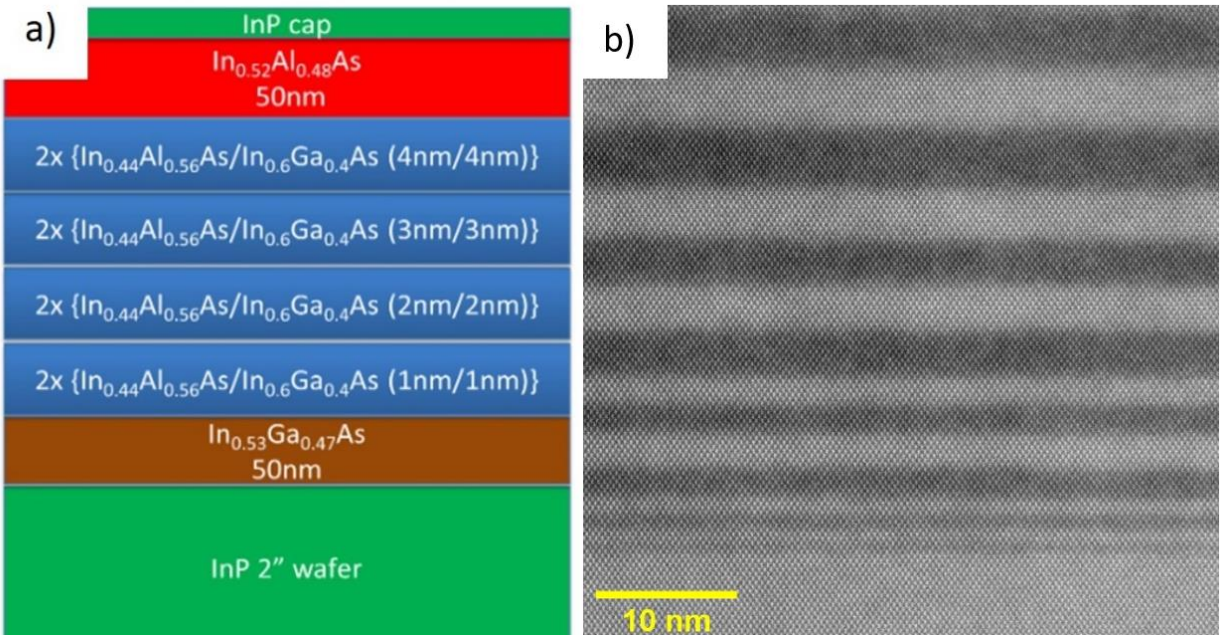


Fig 6.1 (a) Targeted growth for SL structure, **(b)** High-angle annular dark-field scanning transmission electron microscopy **(**HAADF-STEM) image of the grown SL structure.

6.2.2 P17a Structure

The P17a structure, a full 40-stage QCL emitting near 5.0 µm, was grown under the same growth conditions as the thin-SL structure described in Section 6.2.1.

Table 6.2 Growth conditions and targets for two analyzed layers within the P17a structure.

Structure	*Target Composition*	*V/III Ratio*	*Group III Partial Pressure (Torr)*	*Growth Rate (nm/s)*
P17a Barrier	$In_{0.44}Al_{0.56}As$	373	1.43E-3	0.11
P17a Well	$In_{0.57}Ga_{0.43}As$	422	1.27E-3	0.10

The full structure barrier and full structure well targets shown in Table 6.2 are used for the P17a sample. The P17a sample has many layers that contribute to the full QCL structure. However, it also contains layers with the same or similar target thicknesses and compositions and growth conditions as the thin SL sample. TEM images for the P17a sample were acquired, used for the reconstruction calibration, and are shown in Figure 6.2 below.

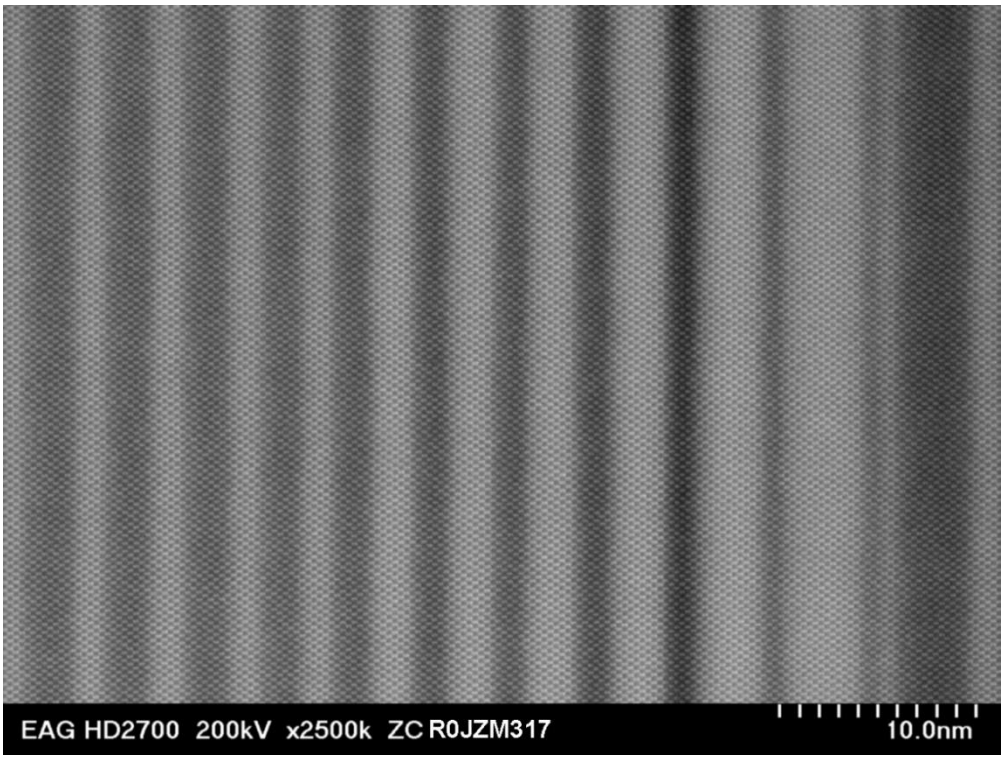


Fig 6.2 STEM high-resolution imaging used to determine the thicknesses of various layers within the P17a sample. Image acquired by EAG Laboratories.

The full-stage thickness is found to be 50.1 nm via TEM, while the growth target was 50.3nm. HR-XRD also showed that the final growth was 0.4% thinner than the target, which is in excellent agreement with the TEM measurement. While TEM provides a clearer picture of individual layers, if calibrations are

being fit to the full QCL structure, then HR-XRD yields a highly accurate result without the intensive process of TEM specimen preparation and analysis.

6.2.3 5g_mod Structure

A full QCL structure emitting at λ~ 4.55 μm was grown on a (001) InP substrate by MOCVD in a close-coupled showerhead (3x2") configuration at a temperature of 605°C. The reactor pressure was held at 100 torr, a 5 sec interruption time was used between layers, and the susceptor was rotated at 100 rpm. The target compositions, V/III ratios, group-III partial pressures, and growth rates are shown in Table 6.3.

Table 6.3 V/III ratios and growth rates for layers grown within the active region.

Layer	*Target Composition*	*V/III Ratio*	*Group III Partial Pressure (Torr)*	*Growth Rate (nm/s)*
Transition Well	$In_{0.69}Ga_{0.31}As$	321.4	1.55E-3	0.120
Second Well	$In_{0.75}Ga_{0.25}As$	256.4	1.94E-3	0.151
1st & 2nd Active Region Barrier	$Al_{0.65}In_{0.35}As$	499.3	9.96E-4	0.081
3rd Active Region Barrier	AlAs	741.1	6.71E-4	0.059

Growth sources included: trimethylindium (TMIn), trimethylgallium (TMGa), trimethylaluminum (TMAl), phosphine, and arsine. An Epison was employed in a feedback loop to control the TMIn concentration.

The structure analyzed is a 40-stage STA-type QCL. Below the active region there is an 0.1 μm-thick $In_{0.53}Ga_{0.47}As$ layer lattice-matched to InP and doped at n=5x10^{16} cm^{-3} that acts as the high-index lower part of waveguide to increase the optical-mode confinement factor, followed by a lower cladding layer consisting of a 2 μm-thick n=2x10^{16} cm^{-3} InP. Directly above the active region there is another 0.1 μm-thick $In_{0.53}Ga_{0.47}As$ layer, doped at n=5x10^{16} cm^{-3}, as the high-index lower part of the waveguide, a 3

µm-thick InP upper cladding layer, doped at n=2x10^{16} cm^{-3}, and finally a 1 µm-thick highly doped InP (2x10^{19} cm^{-3}) layer that acts as the contact layer.

The three interfaces analyzed are labeled in Figure 6.3 along with the corresponding calculated fractions of the leakage current densities at each interface, displayed on the right y-axis, which are influenced by the IFR scattering occurring at each interface.

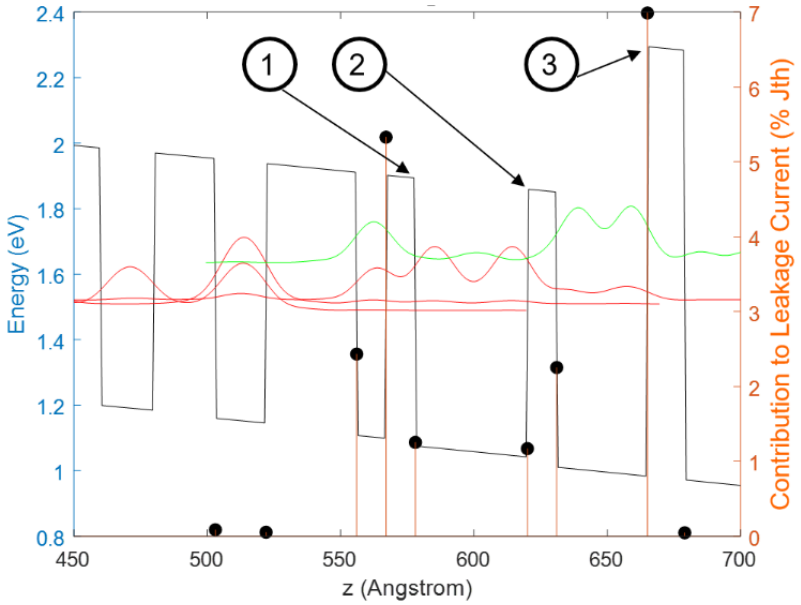


Fig. 6.3 Conduction band energy of the active region of one stage which corresponds to the left y-axis, the percent of threshold-current density that is lost at each interface is shown as black dots which correlate with the right y-axis for the analyzed structure. The red wavy lines represent the wavefunctions corresponding to injector states and the upper laser level, while the green wavy line is the wavefunction of the energy level just above the *ul* level (i.e., the *ul*+1 level). The growth direction is from right to left.

The first analyzed interface is $Al_{0.65}In_{0.35}As$ → $In_{0.69}Ga_{0.31}As$, shown as 1 in Figure 6.3, and the second interface is in inverse order: $In_{0.69}Ga_{0.31}As$ → $Al_{0.65}In_{0.35}As$, shown as 2 in Figure 6.3. These two interfaces were chosen to be analyzed as they bound the thinnest layers in the structure (1.2 nm target

thickness), have the same compositional target, and clad the same InGaAs well, allowing for the observation of differences in roughness parameters when entering or exiting an $Al_{0.65}In_{0.35}As$ barrier. The third interface analyzed is AlAs → $In_{0.75}Ga_{0.25}As$, shown as 3 in Figure 6.3, as it has the highest Al concentration, and thus should have the highest strain. This interface also typically plays the largest role in active-region carrier leakage because it has wavefunction overlap of the *ul* level with the energy state above the *ul* level (i.e., the *ul*+1 level), and the largest conduction band offset.

6.3 Tip Shaping

The samples were prepared using the standard FIB lift-out and sharpening process. The following table shows the progression of the beam currents and the annular-milling diameters. This same process was used for all three samples, however, for the full structures, specimen prep was a bit more challenging than typical because the region of interest is 3-4 micron below the surface, which required relatively large wedges be lifted out in order to assure that the active region was present in the wedge.

Table 6.4 Sequence of lowering FIB milling voltages and currents as a function of the annular mill diameter.

FIB Beam Current	Annular Mill Diameter (μm)
30 kV/50 pA	3
	2
	1.5
30 kV/20 pA	1
	0.6
30 kV/10 pA	0.2
30 kV/2 pA	0.14
2 kV/100 pA	Blanket Mill

A sharpened tip from a full QCL structure looks similar to Figure 6.4 below. It is clear from the image below that there is a significant amount of extra material compared to a typical APT tip. If the upper cladding from a full QCL structure were to be etched off such that the active region was closer to the surface, a smaller wedge could be lifted out and a cleaner and more streamlined process of tip shaping could be achieved.

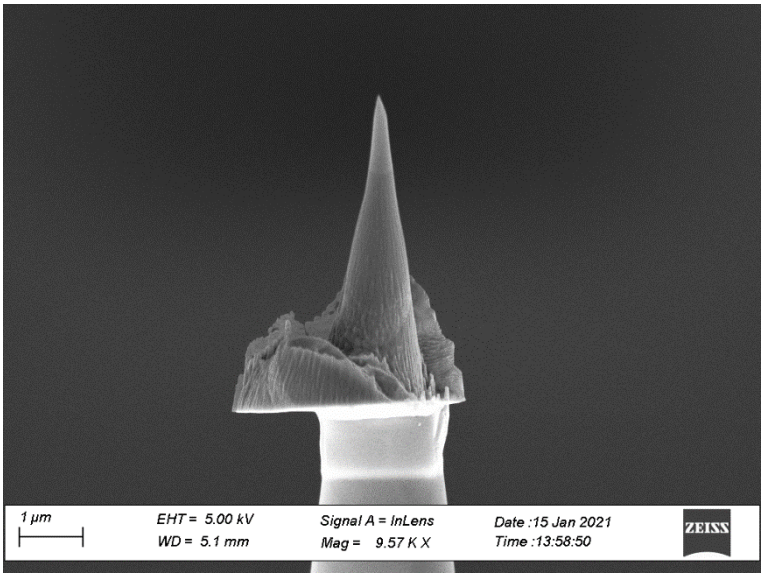


Fig. 6.4: APT tip fabricated with a Ga FIB from a full QCL structure.

6.4 Reconstruction

All three samples contain similar alternating layers of InGaAs and InAlAs and because of this, while there are differences in the reconstructions, all use a shank-angle reconstruction to account for the changing evaporative fields between the two different materials. The differences in reconstruction methods among the different samples is discussed in the following sections.

6.4.1 Thin SL Sample

The thin SL samples were reconstructed in IVAS 3.8.4 using the shank-angle model which avoids artefacts from the voltage drops due to the different evaporation fields between the alternating InAlAs/InGaAs layers. Furthermore, to avoid distortions in the peripheral regions of the reconstruction, only the center of each reconstruction is considered for extracting concentration profiles. The HAADF-TEM images of the thin-SL sample indicated a total thickness of 42 nm excluding the upper and lower cladding layers. The APT reconstruction was calibrated to be within 5% of this measured thickness.

6.4.2 P17a Sample

Like the SL sample, the P17a sample was reconstructed with the shank-angle approximation method and only a smaller region of interest (ROI) from the center of the tip was considered. The TEM analysis of this structure yielded a full stage thickness of 50.1 nm. The APT reconstruction was calibrated such that the full stage thickness was 50.1 nm using the LandMark reconstruction feature.

6.4.3 5g_mod Sample

As with the previous samples, this sample was reconstructed with the shank-angle approximation and used an ROI along the center of the evaporated tip for analysis. The full 40 stage STA-RE QCL structure, with a stage thickness of 45.3 nm, was calibrated with high-resolution x-ray diffraction (HR-XRD) which can be seen in Figure 6.5.

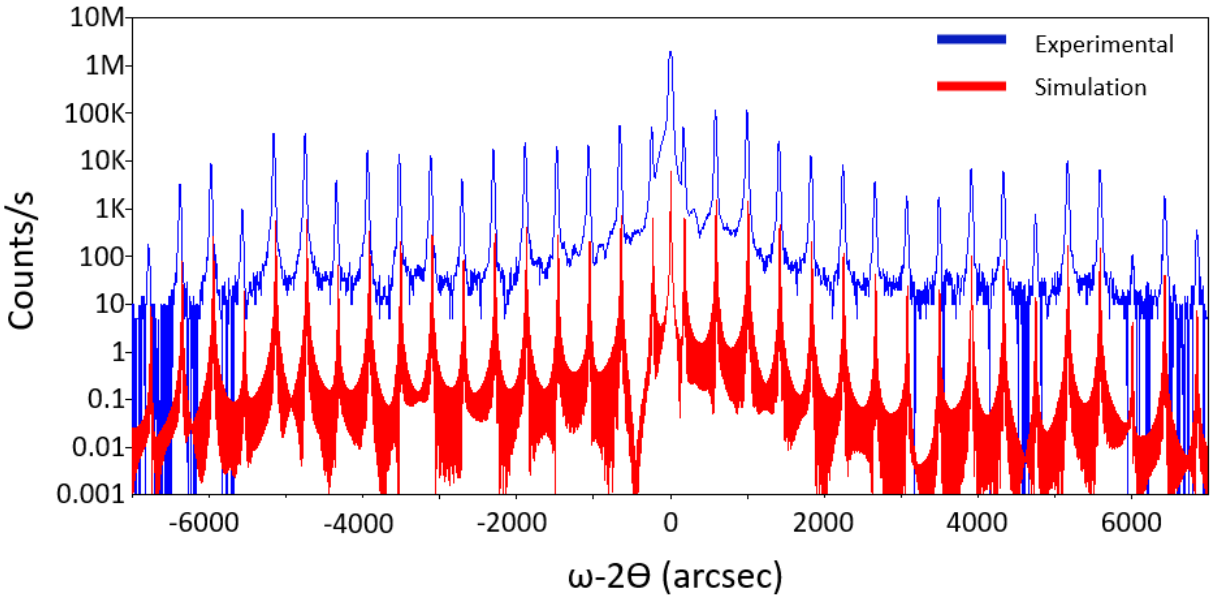


Fig. 6.5 Experimental HR-XRD is shown in blue, and the simulation is shown in red. By matching the simulated HR-XRD to experimental, the thicknesses and compositions of the final structure are extracted.

6.5 APT Results

The extracted values and results for the three different sample groups are separately presented and discussed in the following sections.

6.5.1 Thin SL Sample

The center regions of the two analyzed tips were analyzed to avoid the inclusion of artefacts near the edges of the reconstructions. The first reconstruction of an 18 nm diameter x 35 nm long cylinder in

the center of the specimen is shown below in Figure 6.6 (c), next to the HAADF-STEM (b) image, and a schematic representation of the grown structure (a).

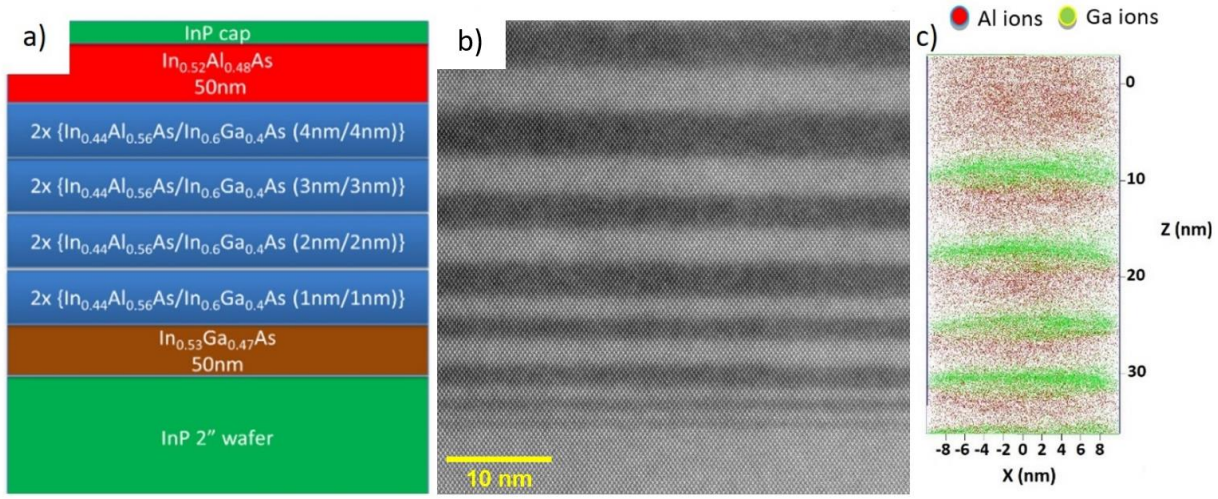


Fig. 6.6 (a) Targeted growth for SL structure; **(b)** HAADF-STEM image of the grown SL structure; **(c)** Region of interest analyzed via APT from the SL.

A 1D concentration profile along the axis of the cylinder and perpendicular to the layers is mapped through this volume to measure the Al and Ga group-III ratio as a function of position. The 1D concentration profile is averaged across the entire ROI with a voxel size of 1 nm x 1nm x 1nm, which leads to lower uncertainty in the results for analysis volumes with a larger diameter. The 1D concentration

profile for the first thin-SL sample is shown in Figure 6.7 below. On the far left of the figure, at 15 nm, the atomic percent Al is in excellent alignment with the target composition of 48% Al.

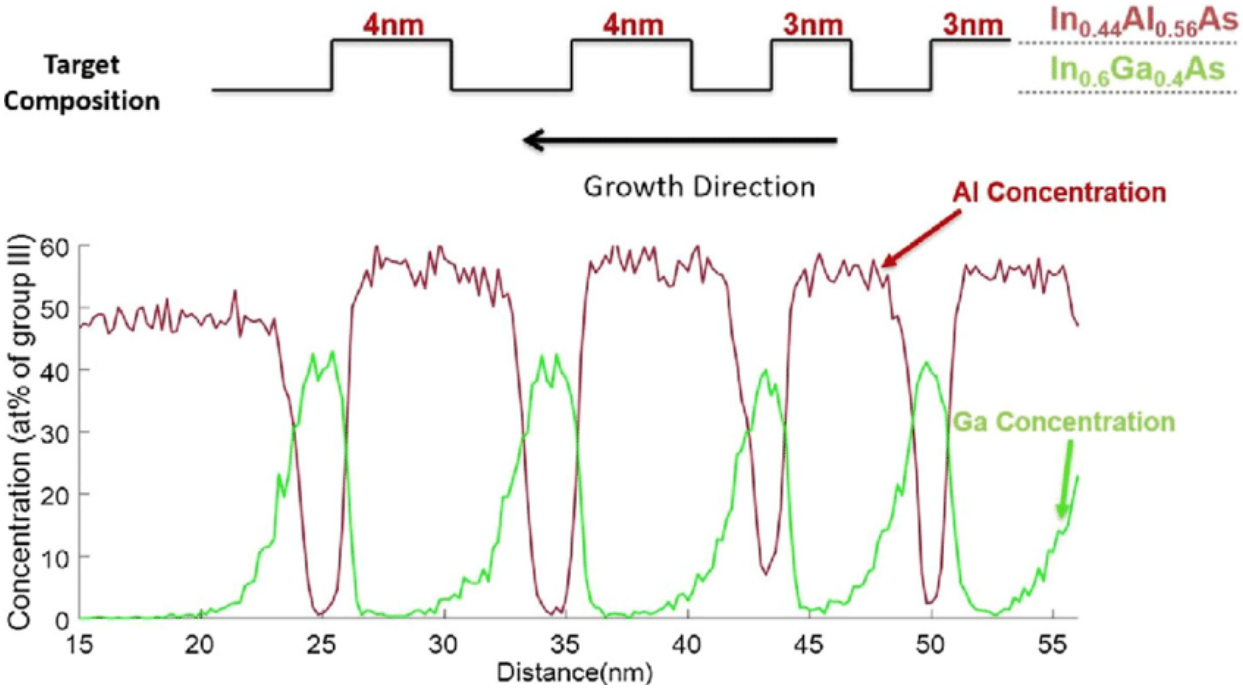


Fig. 6.7 1D concentration profile of the atomic percent of group III for the first tip through the upper layers of the SL with the targeted thicknesses and compositions above.

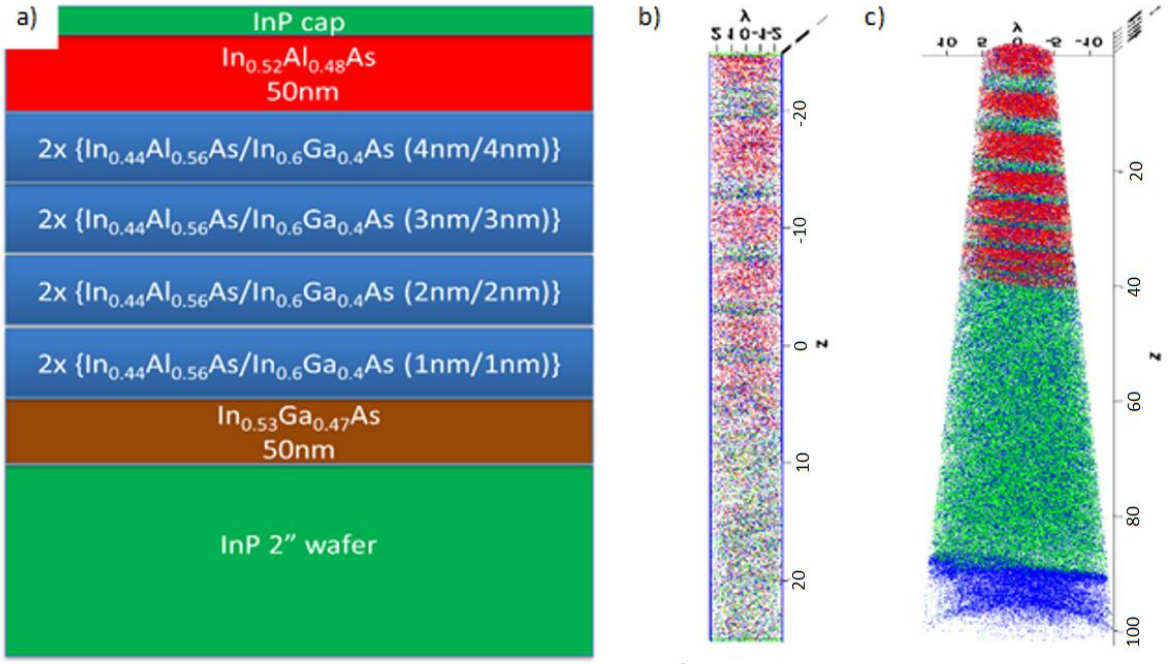


Fig 6.8 (a) target compositions and thicknesses for the SL sample, **(b)** ROI from the reconstruction of the narrow tip, **(c)** full reconstruction of the narrow tip containing the thin SL layers.

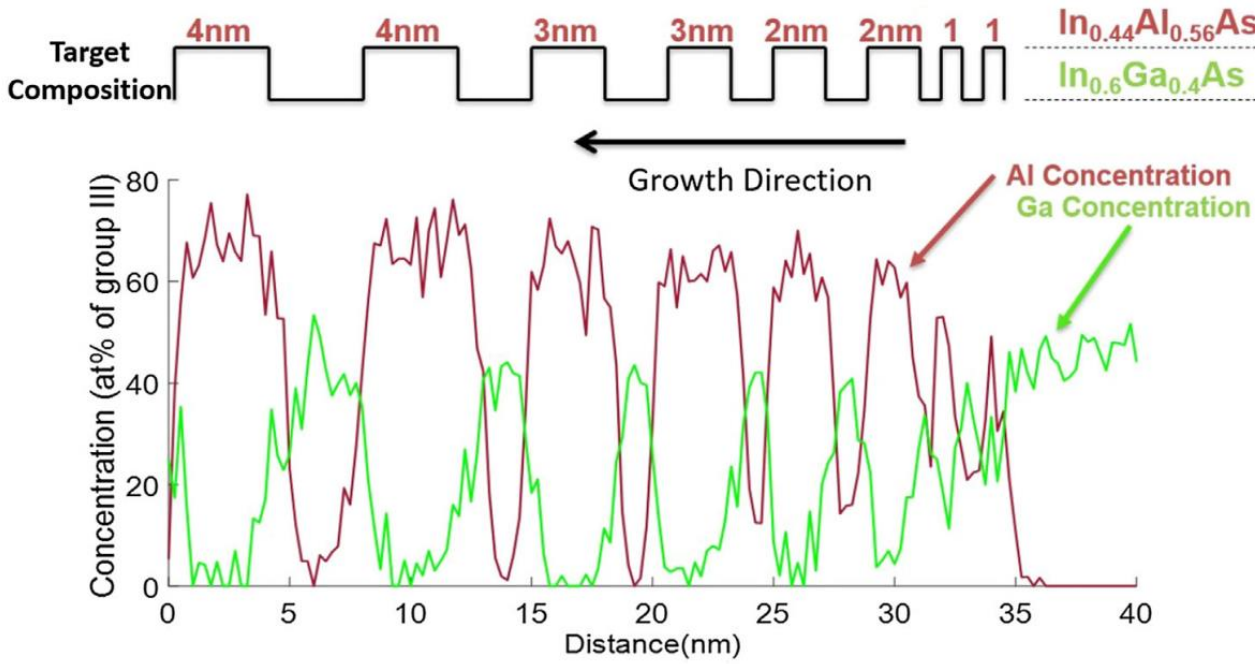


Fig. 6.9 1D concentration profile of the atomic percent of group III for the second tip from the top of the SL through the thin 1 nm-layers with the targeted thicknesses and compositions above.

The thinner tip had a much narrower tip profile which resulted in a significantly reduced volume for analysis, shown in Figure 6.8, resulting in a noisier concentration profile, shown in Figure 6.9. There does appear to be an asymmetry in the concentration profiles across the interfaces, which has been previously observed [4], however, the corresponding indium segregation was not observed in this study. To disambiguate potential artefacts as the source for the asymmetry , the APT data collection would need to be run in the opposite direction, a subject for future studies. Compiling the average composition as a function of the experimentally measured layer thicknesses across both tips is shown in Figure 6.10. As

the layer thickness becomes thinner, the amount of Al or Ga incorporation significantly decreases as well, ~15% deficient at 1 nm for Al, and ~10% deficient at 1nm for Ga with respect to the target compositions.

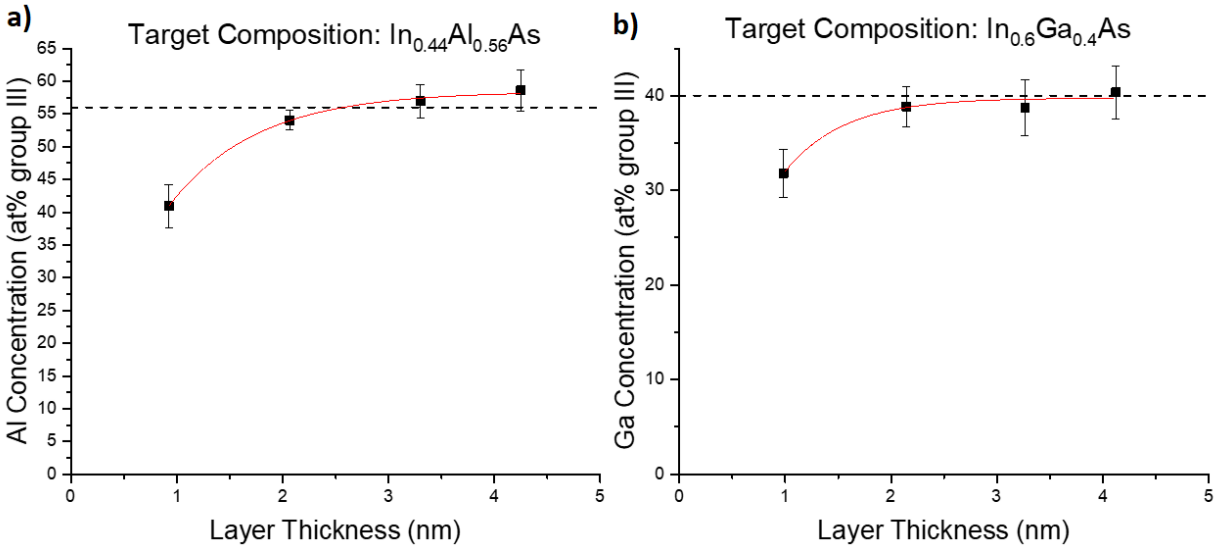


Fig. 6.10 Averaged aluminum **(a)** and gallium **(b)** atomic percent of group III as a function of experimentally measured layer thickness.

6.5.2 P17a Sample

Quantifying with APT the interfacial compositional grading between the first and the last grown stage for the full QCL structure and correlating this with a solid-state diffusion model using the extracted diffusion coefficient described by Rajeev [5], would help elucidate whether significant diffusion occurs during the relatively long growth time of the active-region stages. The tip fabricated from the P17a structure [4] contains the 1st through the 12th stage, shown in Figure 6.11, which, based on the diffusion model used in [5], would be expected to reveal an intermixing length difference of ~0.3 nm, which is within our error for this measurement. Therefore, from this APT measurement we are unable to make a definitive statement regarding the involvement of solid-state diffusion. However, the thin barriers and wells in the QCL can be compared to the two SL tips as they have nearly the same target composition and

the same growth conditions. The measured compositions track well with the previous results, as shown in Figure 6.12.

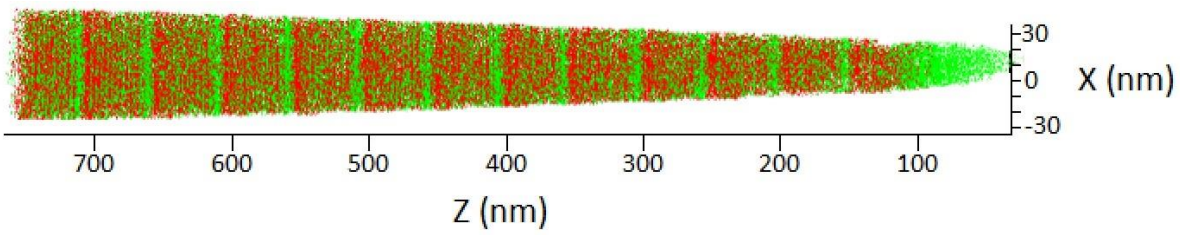


Fig. 6.11 APT reconstruction of the upper 12 stages of the full QCL structure and InGaAs upper waveguide using the Landmark Reconstruction feature in IVAS 3.8.4. The red and green points correspond with Al and Ga atoms, respectively.

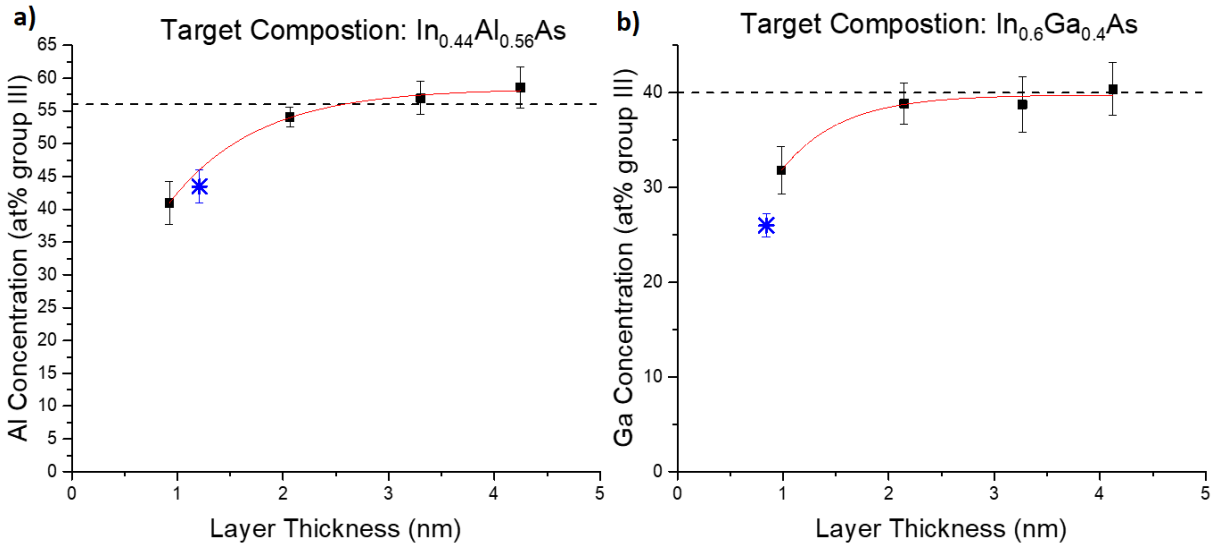


Fig. 6.12 Averaged aluminum **(a)** and gallium **(b)** atomic percent of group III incorporated with the additional data point in blue from the full QCL structure with the layer thicknesses measured via HAADF-STEM. Full QCL targets: $In_{0.44}Al_{0.56}As/In_{0.57}Ga_{0.43}As$.

Previously reported TA-type QCLs [6] emitted at a wavelength (4.8 μm) that was longer than simulated value (i.e., 4.6 μm) using a k*p model and assuming the targeted compositions. APT was performed to determine the actual compositions of the thinnest InAlAs barrier layer at the front of active region within a given stage, as that layer has a large impact on the wavefunctions within the full QCL structure. As can be seen from Figure 6.10 and 6.12, the amount of Al/Ga incorporated in $In_{0.44}Al_{0.56}As/In_{0.6}Ga_{0.4}As$ has a layer-thickness dependence, in that it is significantly decreased for layers thinner than 2 nm. This decrease in Al concentration has been shown to effectively lower the barrier height, resulting in a longer emission wavelength in QCLs [7]. To achieve the modeled QCL emission wavelength, an overshoot for the thinnest InAlAs layer may be necessary [7].

The QCL structure reported in [3] was regrown under the same conditions and a similar STA-type design, except with an overshoot of 5% in the gas phase Al/In ratio to increase the Al content in the thinnest injector barrier layer. This resulted in a modeled and experimental wavelength of 4.6 μm. With this overshoot the exact amount of Al incorporated in these layers is unknown, but it can be assumed it more closely resembles the target composition.

6.5.3 5g_mod Sample

The main purpose behind the analysis of this sample is the extraction of the IFR parameters as outlined by Grange and Mukherjee [8,9].

Previously, a type of interfacial mixing width was extracted from MOCVD grown layers with relevance to QCL active regions and fit to a thermal diffusion profile [5]. However, Grange does not make this assumption and extracts a term, 'L', described in the equation below:

$$c(z) = c_o + d_o \operatorname{erf}\left[\frac{2\sqrt{ln(2)}(z-z_o)}{L}\right] \tag{6.1}$$

where, c_o and d_o are offset and scale parameters, respectively, and z_o is the center of the interface.

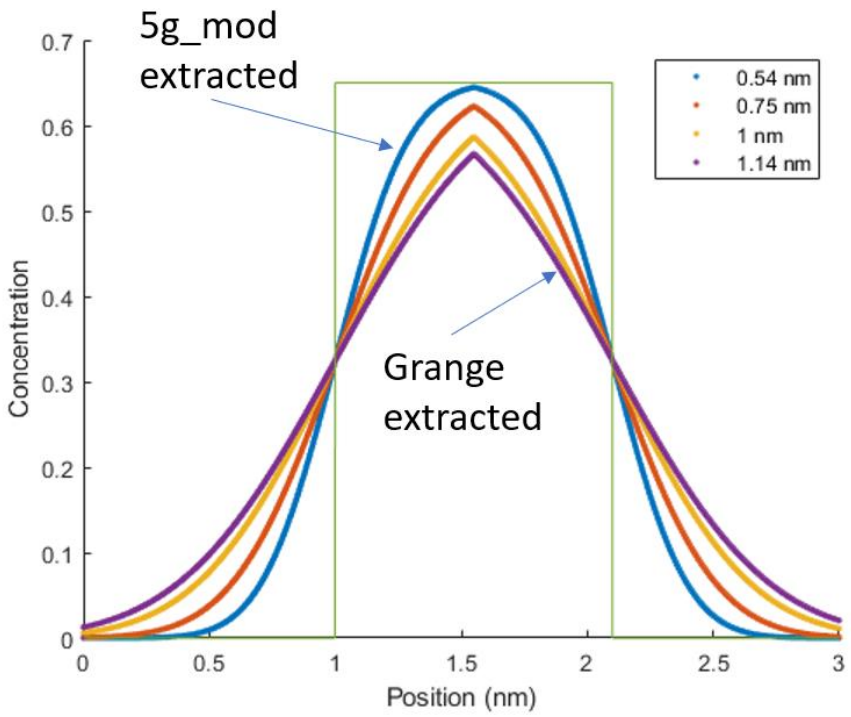


Fig. 6.13 The square function is the idealized concentration profile for a thin barrier (1.1 nm) assuming the concentration of the element of interest goes to zero outside the thin barrier. This is calculated for the value extracted from Grange's paper, the value extracted from the samples presented, and values in-between.

It can be seen from Figure 6.13, if it is assumed that the concentration of the element of interest goes to zero outside of the barrier, then the peak concentration reached has a strong dependence on the extracted mixing width and barrier thickness. However, unfortunately, the layers and materials of interest presented here do not go to zero outside of the target barrier, which can be seen in Figure 6.14 below.

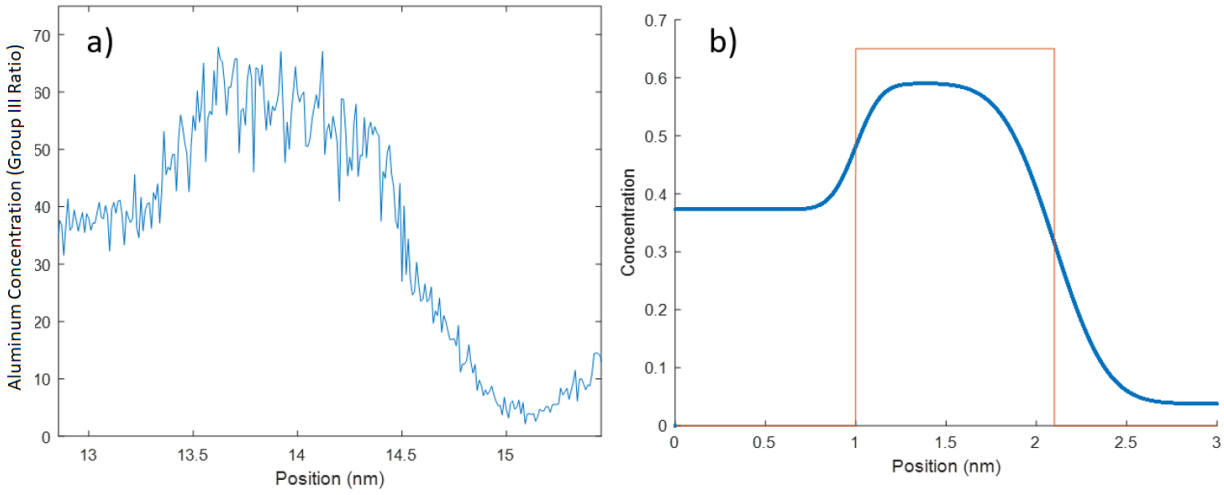


Fig. 6.14 (a) Experimental aluminum concentration profile for a thin barrier (target 1.1 nm) within the 5g_mod structure. **(b)** Aluminum concentration profile for the same thin barrier extracting interfacial mixing widths for the left and right side of the barrier of 0.26 nm and 0.54 nm, respectively. The red square function shows the target thickness and aluminum group III ratio.

From Figure 6.14, it is clear the left and the right side of the barriers and wells may have different interfacial mixing widths, which is why analyzing interfaces 1 and 2 (see Fig, 6.3) is of such importance, as they share a common composition for the adjacent quantum well. This allows the for a direct comparison of the interfacial mixing widths of entering versus exiting a barrier while maintaining common compositional targets and thicknesses. An example of this fitting is shown in Figure 6.15 below. This process is performed across the three different interfaces for three stages from the reconstruction.

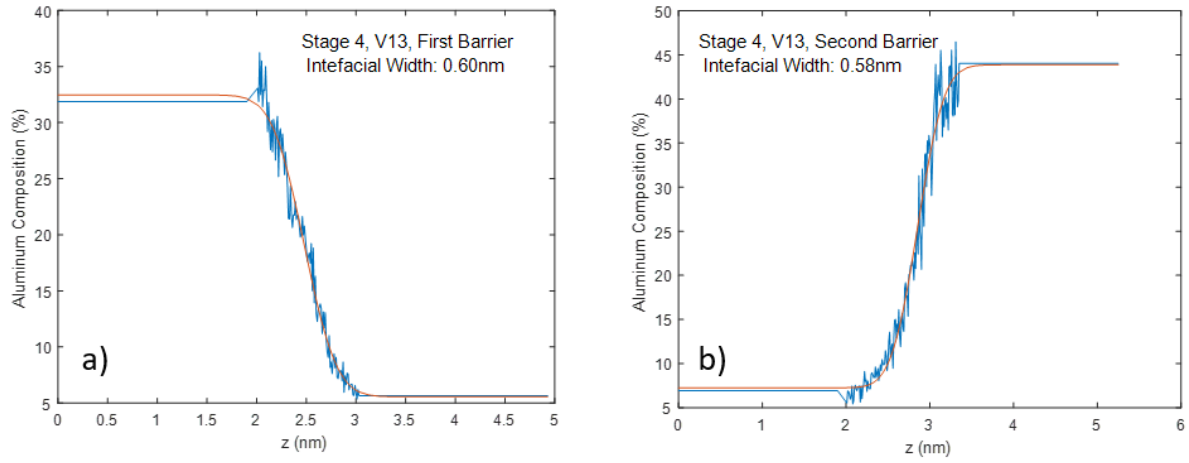


Fig. 6.15 (a) Interface 1 from stage 4 is fit with equation 6.1 and the extracted interfacial width is 0.6 nm. **(b)** Interface 2 from stage 4 is fit with equation 6.1 and the extracted interfacial width is 0.58 nm.

The two key parameters that are extracted however, are the Δ and Λ values, which, as discussed in Chapter 2, are the RMS roughness and correlation length, respectively. These two parameters play a significant role in the IFR scattering rate and, in turn, in the amount of IFR-triggered leakage. The method employed by Mukherjee and Grange [8,9] is applied here as well. The voxel size is 2 nm x 2 nm in the x,y direction, and 0.5 nm in the z-direction. Larger voxel sizes can result in positional errors, but smaller voxel sizes can result in statistical errors. The dimensions of the voxel used here was in line with the user manual for characterizing interfaces in the APT data visualization and analysis software, IVAS. The center of the interface is defined as the loci of points where the Al concentration is one half the difference in aluminum concentration between the two maximum/minimum values in the bounding well and barrier. From previous work [5] and the results here, it is clear that the targeted ternary layers are actually quaternaries of AlInGaAs with varying compositions on the group III site of group III through the various layers. With the specified aluminum concentration, the isoconcentration surface can be defined in IVAS and exported to a BCR file. The BCR file can then be read using an open-source software, Gwyddion. Once in Gwyddion,

the BCR file can be converted to a text file which can be imported by Matlab. An example of one of the extracted isoconcentration surfaces can be seen in Figure 6.16 below.

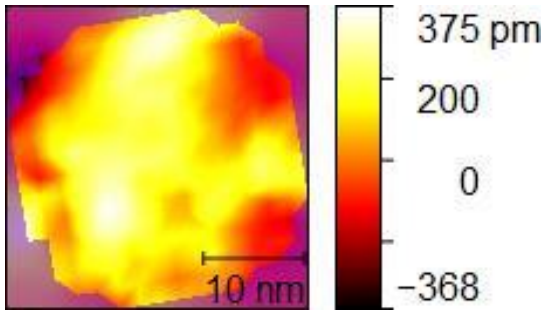


Fig. 6.16 Aluminum isoconcentration surface from the 5g_mod sample.

Once the isoconcentration surface is defined, it is shifted in the z-direction such that the average value is at 0 as shown in the equation below.

$$h(\vec{\rho}) = \vec{\rho}(x,y) - z_0 \tag{6.2}$$

where z_0 is the average position in the z-direction, and $\vec{\rho}(x,y)$ is the z-height as a function of position in the x,y plane. After this calculation is performed, a height-height correlation function (HHCF) can be applied to the isoconcentration surface as shown in the equation below.

$$H(\vec{\tau}) = \langle |h(\vec{\rho}) - h(\vec{\rho} + \vec{\tau})|^2 \rangle_{\vec{\rho}} \tag{6.3}$$

where $\vec{\tau}$ is a vector in the x,y plane. This equation yields a plot that has the characteristic shape as shown in Figure 6.17. At small $\vec{\tau}$ values, there is a large dependence on $\vec{\tau}$, however, as $\vec{\tau}$ increases, this dependence should decrease and potentially vanish. The data output by equation 6.2 can then be fit using the equation below.

$$H_{fit}(\vec{\tau}) = 2\Delta^2\left[1 - e^{-\left(\frac{\tau}{\Lambda_{||}}\right)^2}\right] \tag{6.4}$$

where Δ is the amplitude of the roughness from the isoconcentration surface and $\Lambda_{||}$ is the in-plane correlation length.

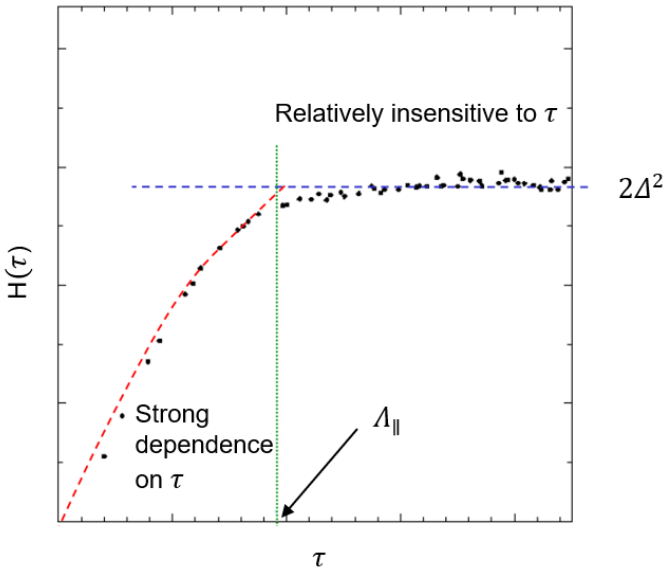


Fig. 6.17 Schematic of a HHCF showing how at small $\vec{\tau}$ values there is a strong dependence on $\vec{\tau}$, however, at large $\vec{\tau}$ values, it is relatively insensitive to $\vec{\tau}$.

With the process in place, the three specified interfaces are analyzed and the in-plane IFR parameters are extracted. A characteristic extracted HHCF for each interface is shown in the figures below.

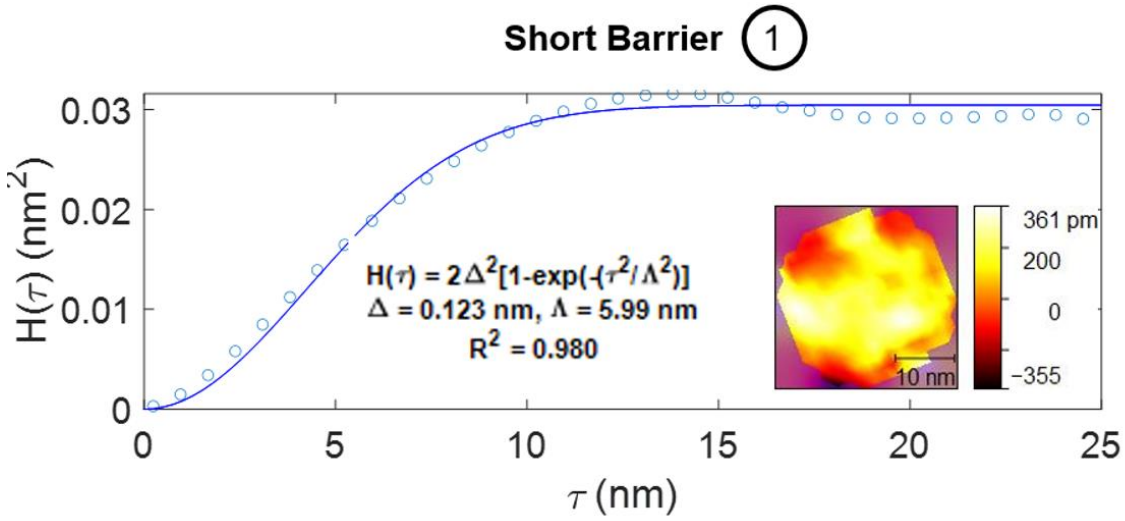


Fig. 6.18 A characteristic HHCF from the first barrier labeled as interface #1. The blue open circles are the experimental data and the solid blue line is the fit to the data. The equation for the fit and the extracted in-plane IFR parameters are included.

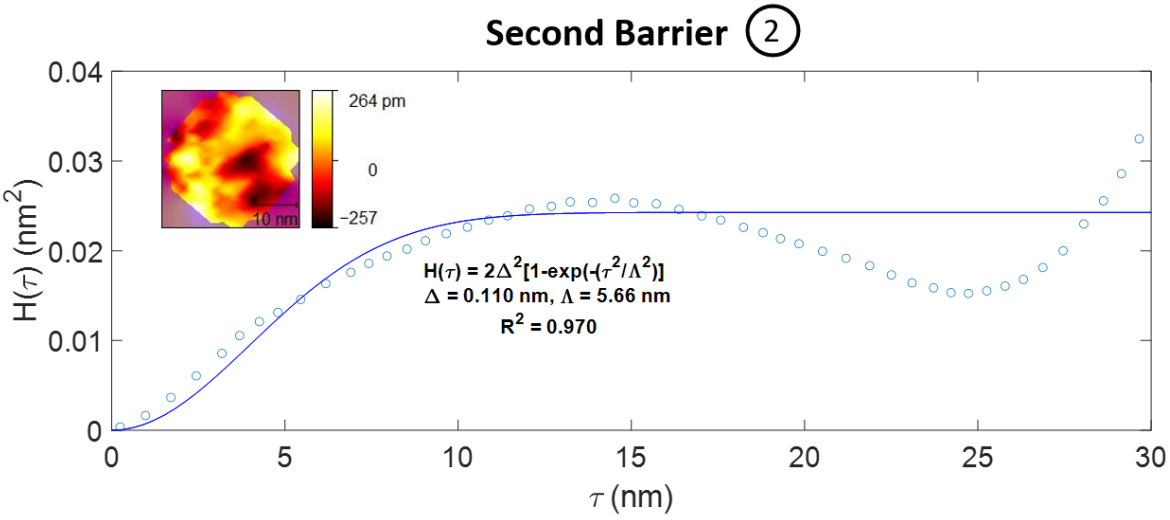


Fig. 6.19 A characteristic HHCF from the second barrier labeled as interface #2. The blue open circles are the experimental data and the solid blue line is the fit to the data. The equation for the fit and the extracted in-plane IFR parameters are included.

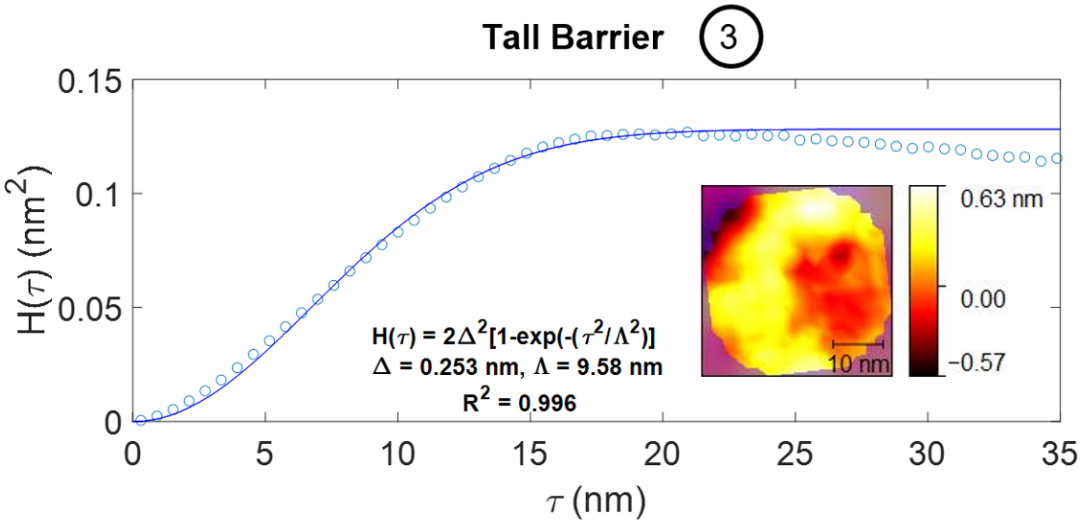


Fig. 6.20 A characteristic HHCF from the third barrier labeled as interface 3. The blue open circles are the experimental data and the solid blue line is the fit to the data. The equation for the fit and the extracted in-plane IFR parameters are included.

As these reconstructions did not use the LandMark Reconstruction feature in IVAS, there was a concern that bowing of the isoconcentration surfaces which may lead to incorrect extracted parameters. To analyze the impact of bowed surfaces, a Matlab script was written to create a 3D plane with specified roughness parameters. The plane is then analyzed with the HHCF script used to analyze the isoconcentration surfaces output by IVAS to determine the sensitivity of the extracted values on the reconstruction. A 3D plane with Δ=0.1 nm and Λ=1 nm was constructed and can be seen in Figure 6.21.

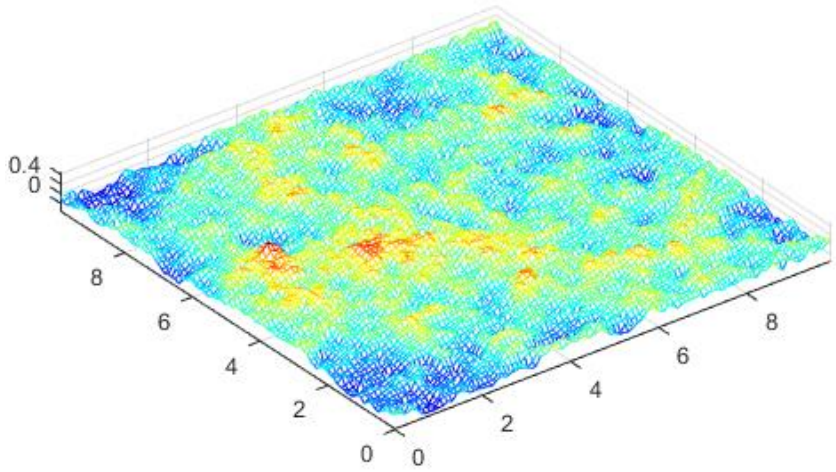


Fig. 6.21 Arbitrary 3D plane with nm as the units in all 3 directions. The specified Δ and Λ IFR-parameter values are 0.1 nm and 1 nm, respectively.

Using the surface shown in Fig. 6.21 the HHCF data was generated and fit with equation 6.4, as done above. The extracted in-plane IFR parameters are Δ=0.1 nm and Λ=0.8 nm showing excellent agreement for the amplitude of roughness and a correlation length; that is, only about 20% less than specified for Λ. As it can be seen from Fig. 6.22 below, there is not a strong transition between a strong and weak dependence on $\vec{\tau}$ which is most likely why the correlation length is lower than expected. This is the extracted values for a perfectly planar interface to use as a baseline. With the baseline in place,

bowing is introduced to the surface to determine the resulting effect on the extracted roughness parameters.

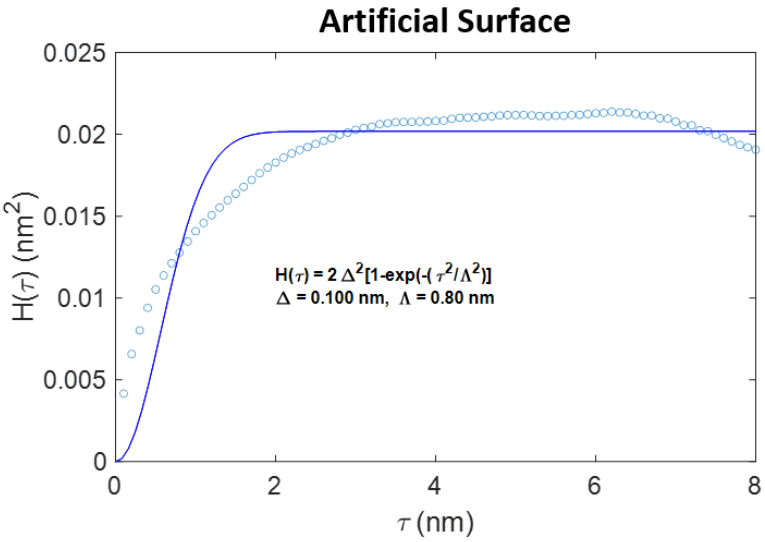


Fig. 6.22 The HHCF results from the artificial surface shown in Fig. 6.21 and the resulting fit and extracted in-plane IFR parameters. The open blue circles are the calculated values from the surface and the solid blue line the fit to the experimental data.

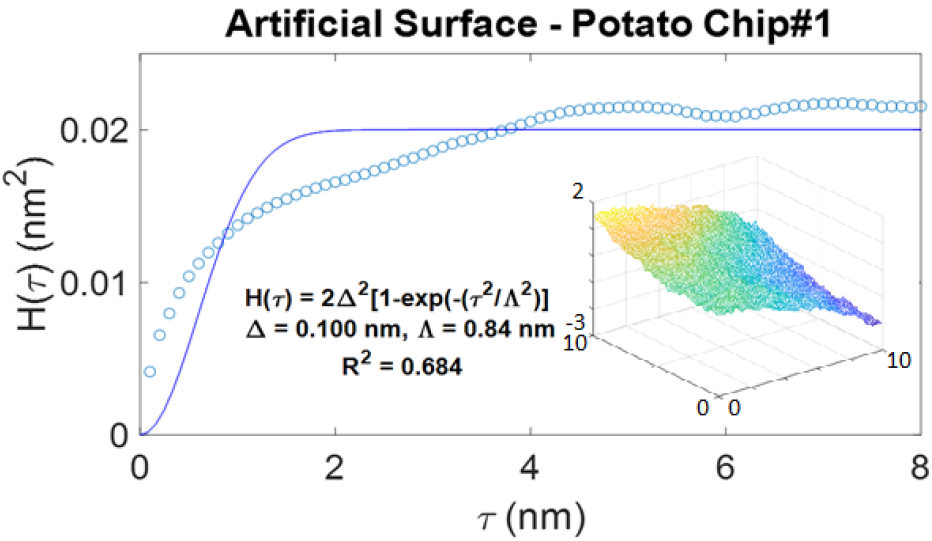


Fig. 6.23 The HHCF results from the artificial surface, shown in the inset (axes is nm), and the resulting fit and extracted in-plane IFR parameters with bowing introduced. The open blue circles are the calculated values from the surface and the solid blue line the fit to the experimental data.

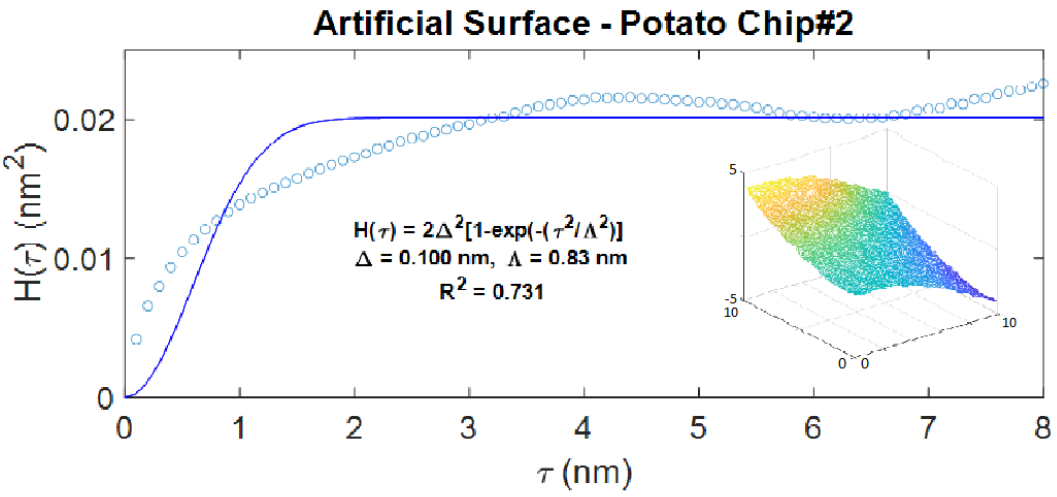


Fig. 6.24 The HHCF results from the artificial surface, shown in the inset (axes in nm), and the resulting fit and extracted in-plane IFR parameters with bowing introduced. The open blue circles are the calculated values from the surface and the solid blue line the fit to the experimental data.

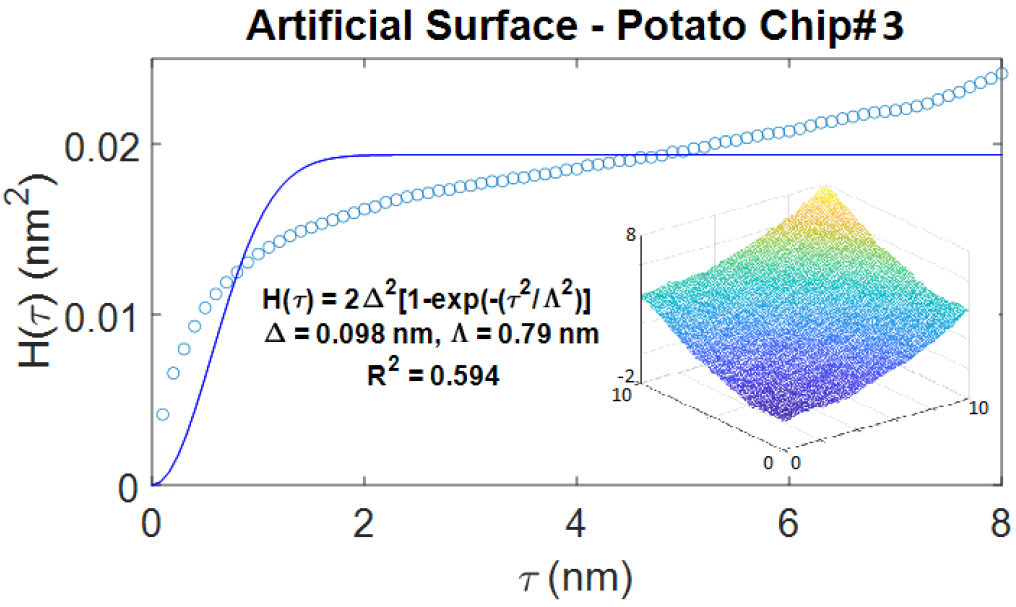


Fig. 6.25 The HHCF results from the artificial surface, shown in the inset (axes in nm), and the resulting fit and extracted in-plane IFR parameters with bowing introduced. The open blue circles are the calculated values from the surface and the solid blue line the fit to the experimental data.

Even with the inclusion of significant bowing and non-planarity of the interface, upwards of 10 nm off-axis, the extracted Δ and Λ values are within 2% and 5%, respectively, which can be seen in Figures 6.23-6.25. This shows that slight variations in the planarity of the reconstructed interface should not have a significant impact on the extracted IFR parameters. Despite this, planarity of the interfaces was a goal of the reconstructions and are planar to the best of our ability without the use of the LandMark Reconstruction feature in IVAS. With this knowledge, the IFR parameters from the isoconcentrations from the 5g_mod sample are extracted and averaged across three stages of the reconstruction. The averaged values for the three interfaces of interest are shown in the table below. The script used is the same as the one used to generate figures 6.18-20 and extract the in-plane IFR parameters.

Table 6.5 Extracted IFR parameters and interfacial mixing width from the three interfaces of interest averaged across three stages. The strain is calculated at the mid-point of the change in Al concentration through the interface of interest. The column labeled Differential Strain Relative to InP is the magnitude of the strain between the compressive-strained barrier and tensile-trained well relative to InP.

Interface	*Δ (nm)*	*Λ (nm)*	*Interfacial Mixing Width (nm)*	*Differential Strain Relative to InP (%)*
$In_{0.69}Ga_{0.31}As \rightarrow Al_{0.65}In_{0.35}As$	0.145 (±0.02)	6.61 (±0.76)	0.54 (±0.05)	2.3
$Al_{0.65}In_{0.35}As \rightarrow In_{0.69}Ga_{0.31}As$	0.120 (±0.01)	5.75 (±0.36)	0.61 (±0.08)	2.3
$AlAs \rightarrow In_{0.75}Ga_{0.25}As$	0.201 (±0.02)	7.08 (±1.06)	0.68 (±0.2)	5.1

From Table 6.5, the first two interfaces have very similar Δ and Λ values, showing that entering versus exiting a barrier or well for layers that share the same level of strain or have the same target thickness and composition appears to not play a significant role in interface roughness. Looking at the difference between an upstream and downstream barrier also does not appear to show any relation with the interfacial mixing width; even the tall barrier, or third interface, shows similar interfacial mixing width values when including the error bars. Comparing the AlAs → $In_{0.69}Ga_{0.31}As$ interface with the other two interfaces a large discrepancy in the extracted Δ values can be seen; roughly a 50% increase in the RMS roughness amplitude, Δ. We can also see that the strain at the interface, defined to be the sum of the strain between the two layers of interest relative to InP, increases by about 100%, thus showing a potential correlation with the increase in the Δ value.

The IFR scattering rate is proportional to Δ^2, Λ^2, and δU^2 as shown in Equation 2.6 in Chapter 2. A 50% increase in the Δ value, as measured here, would result in a scattering rate more than two times higher than the previous value for a particular interface. That is, the IFR scattering rates at different interfaces appears to correlate well with the amount of strain at each interface. Variations in the roughness values for different interfaces will have a significant impact on the IFR scattering rates and leakage currents for these devices; thus, for an accurate model one should consider the varying levels of strain at interfaces throughout a full QCL active-region structure.

However, as mentioned in Chapter 5, there is also an axial correlation component. The calculation is performed through an interface. Now instead of only considering the isoconcentration at the half-way point in difference in aluminum concentration and extracting the in-plane IFR parameters, the isoconcentrations around the center are considered and how the correlation between subsequent isoconcentrations evolve through the interface is analyzed. Using the following series of equations, the axial correlation length can be extracted from isoconcentrations through an interface.

$$w_{ij}(\vec{\rho}) = h_i(\vec{\rho}) - h_j(\vec{\rho}) \tag{6.5.}$$

$h_i(\vec{\rho})$ is the isoconcentration profile for a concentration designated by 'i' and the same can be said for $h_j(\vec{\rho})$. The difference between these two isoconcentration surfaces is defined as a differential map, $w_{ij}(\vec{\rho})$.

$$\Delta_{ij} = \sqrt{\langle w_{ij}^2(\vec{\rho})\rangle} \tag{6.6}$$

Δ_{ij} is then the roughness of the differential map.

$$\Delta_{ij}^2 = 2\Delta^2\left[1 - C_{\perp}\left(\left|z_i - z_j\right|\right)\right] \tag{6.7}$$

Δ is the rms roughness previously extracted and is assumed to be constant through the interface. z_i and z_j refer to the position of each respective isoconcentration relative to each other. The value used for these two positions is the centroid of the isoconcentration directly output from IVAS. An example of the isoconcentrations as a function of position through an interface is shown in Figure 6.26 below. The points show a clear progression through the interface as a function of aluminum (i.e., group-III) ratio.

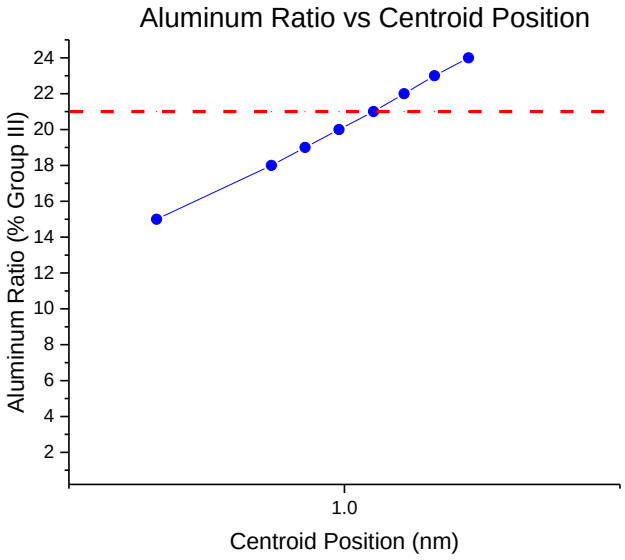


Fig. 6.26 Identified isoconcentrations through the second barrier (interface 2) from the third stage and their corresponding centroid position as blue dots. The dashed red line is the center of the interface.

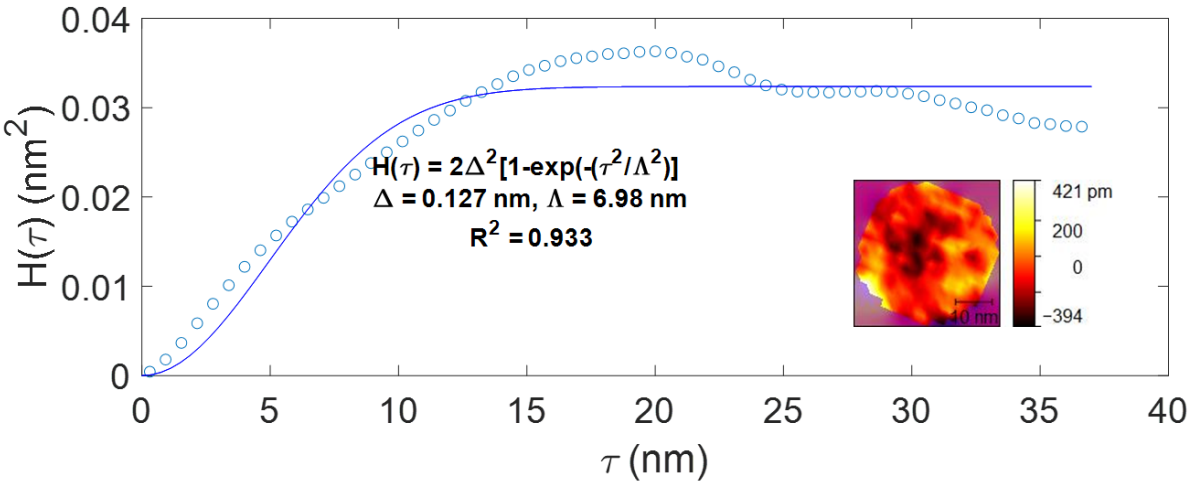


Fig. 6.27 The HHCF for the center of the interface for the second barrier (interface 2) from the third stage. The isoconcentration at the intersection of the red line and the blue dots in Figure 6.26 yields this HHCF.

$$C_{\perp}(\Delta z_{ij}) = 1 - \frac{\Delta_{ij}^2}{2\Delta^2} \tag{6.8}$$

Plotting the data with equation 6.7 should result in a graph with an exponential decay relative to the change in z position. That graph can then be fit with the following equation in order to extract the axial correlation length, $\Lambda_{\perp}$.

$$e^{-\Delta z/\Lambda_{\perp}} \tag{6.9}$$

Performing this calculation for a series of isoconcentrations through an interface for the second barrier (interface 2) in the third stage results in a graph and fit as shown in Figure 6.28.

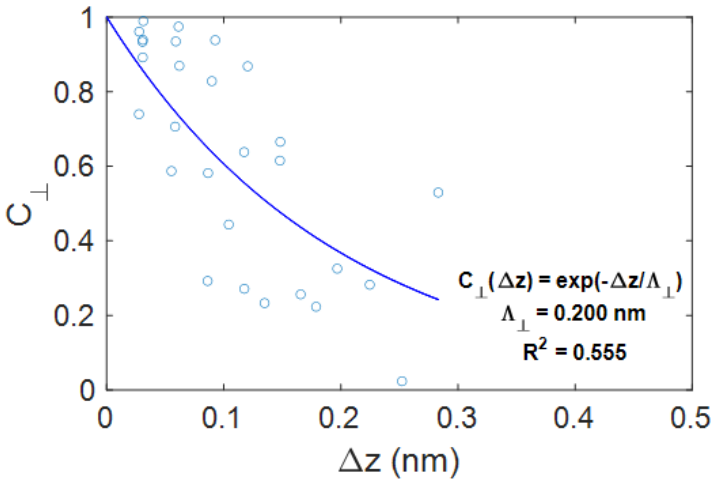


Fig. 6.28 Extraction of axial correlation length from the second barrier (interface 2) from the third stage.

Clearly, the fit to this data does not show a strong correlation. Recent papers by Grange and Mukherjee [8,9] are built on the premise of uncorrelated isoconcentrations within a diffuse interface where the extracted axial correlation length is less than the interfacial mixing width. From the images of the isoconcentrations shown in Figure 6.29, which correlate with the isoconcentrations through the interface shown Figure 6.26, there is some level of correlation. However, like Grange and Mukherjee, it can be seen here, despite the poor fit, the extracted axial correlation length (0.2 nm) is significantly smaller than the extracted interfacial mixing width (0.55 nm) supporting previous findings and allowing the implementation of the axial correlation into a scattering model.

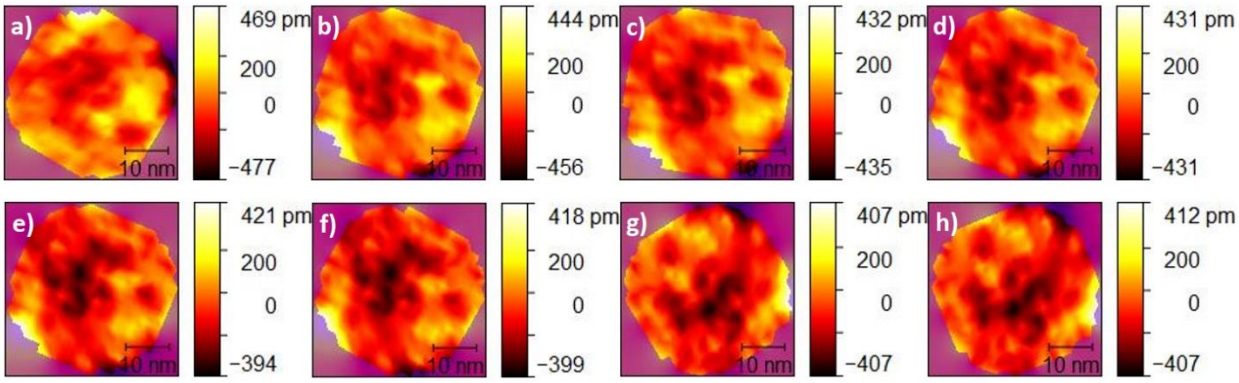


Fig. 6.29 Isoconcentrations through the interface shown in Figure 6.26 and the isoconcentrations used in the calculation for Figure 6.28. **a)** 15% aluminum group III ratio, **b)** 18% aluminum group III ratio, **c)** 19% aluminum group III ratio, **d)** 20% aluminum group III ratio, **e)** 21% aluminum group III ratio, **f)** 22% aluminum group III ratio, **g)** 23% aluminum group III ratio, and **h)** 24% aluminum group III ratio.

6.6 Scattering Model Incorporation

Previous findings [7,9] have shown that interfaces are not atomically flat, however, from the standpoint of IFR scattering and IFR-triggered leakage currents, a formalism has not been developed to date to account for these diffuse interfaces. A formalism has been created and is discussed in detail in Chapter 5. However, what is not known is how the introduction of diffuse interfaces, in the form of graded-interfaces IFR leakage, lattice constant, conduction band edge, and AD scattering, will affect the various lifetime values obtained when using a scattering model for QCL structures.

For continuity, the design that is analyzed is the same as for the sample 5g_mod in the previous section, as the values extracted from APT may vary by design. The incorporation of these changes is done stepwise to see the impact of each. First is the incorporation of quaternary-alloy AD scattering in the graded-interface regions. The AD scattering rate is calculated, at each point throughout the active region, in accordance with Equation 2.6 in Chapter 2, which is then integrated over the entire active region. The change from sharp interfaces and ternary-alloy scattering to graded interfaces and quaternary-alloy

scattering results in somewhat longer lifetimes. However, the *ul* and *ll* lifetimes do not increased proportionally. The *ul* global AD lifetime increases negligibly (i.e., from 1.93 ps to only 2.02 ps), while the *ll* global AD lifetime increases moderately: from 4.41 ps to 6.07 ps (see Table 6.6). This can be understood in that AD scattering is a bulk scattering mechanism unaffected by interface roughness. Thus, unlike IFR scattering, which is expected to strongly be suppressed in the presence of graded interfaces [8], AD scattering is affected only by small changes in conduction-band edge.

The incorporation of the axial correlation length decreases the IFR scattering rate by the term F which is controlled by the dimensionless parameter $L/\Lambda_{\perp}$, discussed in Chapter 5. The scattering rated extracted by Grange for the Si/SiGe interfaces probed was reduced by nearly 65% when compared to the previous method of calculation. Presented here, using the same methods, the IFR scattering rate is reduced by nearly 70% as the interfacial width is approximately 0.5 nm and the $\Lambda_{\perp}$ is approximately 0.1 nm. Reduction in scattering rate results in longer IFR lifetimes, however, here the in-plane IFR parameters, Δ and Λ_{II}, are being changed from 0.13 nm and 15 nm to a range of values for Δ of 0.13 nm to 0.2 nm and for Λ_{II} from 6.2 nm to 7.1 nm from the APT extracted values based off a linear interpolation of barrier height. The *ul* global IFR lifetime decreases from 1.86E5 ps to 17.07 ps, but the *ll* global IFR lifetime increases from 0.224 ps to 0.404 ps. The latter agrees with what is expected from theory [8], while the former is due to the fact that the front two barriers (in the active region) are relatively highly strained, unlike the P1A-design STA-QCL structure [10].

However, surprisingly, the largest change comes from the LO scattering calculated for graded interfaces. The *ul* global LO lifetime decreases from 1.22 ps to 0.765 ps and the *ll* global LO lifetime increases from 0.286 ps to 0.495 ps. This change in lifetimes leads to a significant decrease in LO transition efficiency from 79.1% down to 55.9%.

Table 6.6 Compiled *ul* and *ll* lifetimes comparing sharp interfaces and the introduction of graded interfaces and variable IFR parameters

	5g_mod – Sharp interface	5g_mod – Graded interfaces, variable IFR parameters
$\tau^{AD}{}_{ul,global}$	1.93 ps	2.02 ps
$\tau^{AD}{}_{ll,global}$	4.41 ps	6.07 ps
$\tau^{IFR}{}_{ul,global}$	186015 ps	17.07 ps
$\tau^{IFR}{}_{ll,global}$	0.224 ps	0.404 ps
$\tau^{LO}{}_{ul,global}$	1.22 ps	0.765 ps
$\tau^{LO}{}_{ll,global}$	0.286 ps	0.495 ps

Incorporation of graded interfaces, an axial correlation length of 0.1 nm, variable IFR parameters, and quaternary AD scattering, while significantly changing lifetimes, supports previous results showing that LO-phonon and AD scattering dominate the *ul* global lifetime [10]. There is also agreement with results from sharp-interface, highly-strained STA-QCLs (i.e., the 5g_mod structure) in that LO-phonon and IFR scattering determine the *ll* global lifetimes.

An interesting result of these calculations is the IFR leakage reduced significantly down to ~1% of the threshold current density is being leaked shown in Figure 6.30.

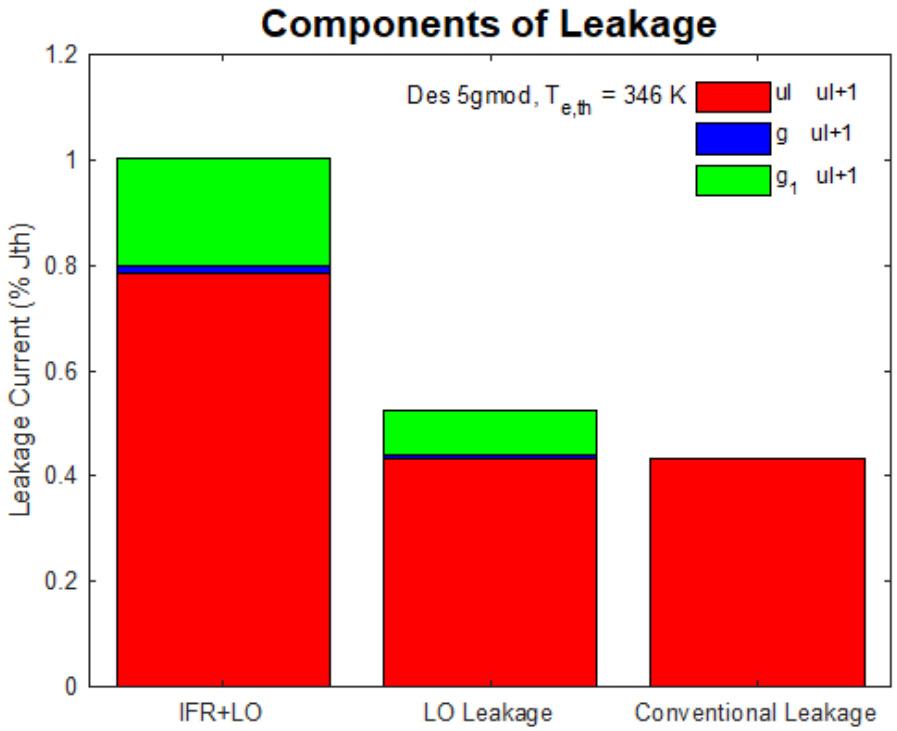


Figure 6.30 Components of leakage for 5g_mod when using an axial correlation length of 0.1 nm, an interfacial mixing width of 0.5 nm, Δ and Λ_{II} are being varied from 0.13 nm to 0.2 nm and 6.2 nm to 7.1 nm, respectively, from the APT extracted values, and quaternary AD scattering.

However, the leakage currents have an exponential dependence on the electronic temperatures and, as can be seen from Figure 6.30, the electronic temperature calculated using Vitiello et al's [11] electron-lattice coupling constant is 346 K. Recent work has shown that electronic temperatures, especially that of the *ul* level, are significantly higher than previously thought [3,10,11]. That is, the nonequilibrium Green's function [NEGF] modeling approach has been used to calculate the transport for a QCL structure emitting at 8.5 µm which yielded an *ul* temperature of 512 K [12]. At shorter wavelengths,

the electronic temperatures, also calculated using NEGF, are found to be even higher [13]. Recent results on 4.9 µm-emitting TA-QCLs [14] have yielded an *ul* electronic temperature of 912 K. The electronic temperatures do increase for the injector states as well: g_3 = 806 K, g_2 = 722 K, g_1 = 587 K, g_0 = 455 K. Furthermore, a recent design of 4.6 µm-emitting STA-QCLs [13], with projected performances higher than those obtained from TA-QCLs [14] and using NEGF modeling and graded interfaces, give an *ul* electronic temperature of 1015 K . Given that, as seen in Figure 6.30, the majority of the leakage is coming from *ul* level and that tripling the *ul*-level temperature results in about an order of magnitude increase in the relative carrier-leakage value, we expect that the leakage current will increase back to ~ 10 %; that is, back to values similar to those found for sharp-interface STA QCLs of conventional *ul*-level electronic temperatures [10].

Overall, it has been found that the lifetimes global *ul* lifetimes decrease and the *ll* lifetimes increase leading to poorer performing devices. Introduction of the axial correlation length leads to significantly reduced leakage currents. However, using the elevated electronic temperatures calculated by using the more accurate NEGF modeling approach, holds potential to counteract the carrier-leakage decrease.

References

[1] B. Gault, A. J. Breen, Y. Chang, J. He, E. A. Jagle, P. Kontis, P. Kurnsteiner, A. Kwiatkowski da Silva, S. K. Makineni, and I. Mouton, "Interfaces and defect composition at the near-atomic scale through atom probe tomography investigations" *J. of Matl. Research*, 33(23), 4018-4030 (2018)

[2] Y. Yu, C. Zhou, S. Zhang, M. Zhu, M. Wuttig, C. Scheu, D. Raabe, G. J. Snyder, B. Gault, and O. Cojocaru-Miredin, "Revealing nano-chemistry at lattice defects in thermoelectric materials using atom probe tomography", *Materials Today*, Vol 32, 260-274, (2020)

[3] D. Botez, J. D. Kirch, C. Boyle, K. M. Oresick, C. Sigler, H. Kim, B. B. Knipfer, J. H. Ryu, D. Lindberg, T. Earles, L. J. Mawst, and Y. V. Flores. "High-efficiency, high-power mid-infrared quantum cascade lasers [Invited], *Opt. Mater. Express*. Vol. 8, No. 5, 1378, (2018)

[4] C. A. Wang, B. Schwarz, D. F. Siriani, M. K. Connors, L. J. Missaggia, D. R. Calawa, D. McNulty, A. Akey, M. C. Zheng, J. P. Donnelly, T. S. Mansuripur, and F. Capasso, "Sensitivity of heterointerfaces on emission wavelength of quantum cascade lasers", *Journal of Crystal Growth*, Vol. 464, 215-220, (2017)

[5] A. Rajeev, C. Weixin, J. D. Kirch, S. E. Babcock, T. F. Kuech, T. Earles, and L. J. Mawst, "Interfacial Mixing Analysis for Strained Layer Superlattices by Atom Probe Tomography", *Crystals* **2018**, 8, 437

[6] J. D. Kirch, J. C. Shin, C. -C. Chang, L. J. Mawst, D. Botez, and T. Earles, "Tapered active-region quantum cascade lasers (λ=4.8 μm) for virtual suppression of carrier-leakage currents", Vol. 48, Iss. 4, 234-235, (2012).

[7] K. Pantzas, G. Beaudoin, G. Patriarche, L. Largeau, O. Mauguin, G. Pegolotti, A. Vasanelli, A. Calvar, M. Amanti, and C. Sirtori, "Sub-nanometrically resolved chemical mappings of quantum-cascade laser active regions", *Semiconductor Science and Technology*, **31**, 055017, (2016)

[8] T. Grange, S. Mukherjee, M. Montanari, L. Persichetti, L. Di Gaspare, S. Birner, A. Attiaoui, O. Moutanabbir, M. Virgilio, and M. De Seta, "Atomic-Scale Insights into Semiconductor Heterostructures: From Experimental Three-Dimensional Analysis of the Interface to a Generalized Theory of Interfacial Roughness Scattering", *Phys*. *Rev. Appl.*, **13**, 044062 (2020)

[9] S. Mukherjee, A. Attiaoui, M. Bauer, O. Moutanabbir, "3D Atomic Mapping of Interfacial Roughness and Its Spatial Correlation Length in Sub-10 nm Superlattices", *Appl. Mater. Interfaces* 2020, **12**, 1728-1739

[10] C. Boyle, K. M. Oresick, J. D. Kirch, Y. V. Flores, L. J. Mawst, and D. Botez, "Carrier Leakage via Interface-Roughness Scattering Bridges Gap Between Theoretical and Experimental Internal Efficiencies of Quantum Cascade Lasers", Appl. Phys. Lett. **117**, 109901 (2020).

[11] M. S. Vitiello, T. Gresch, A. Lops, V. Spagnolo, G. Scarmarcio, N. Hoyler, M. Giovannini, and J. Faist, "Influence of InAs, AlAs δ layers on the optical, electronic, and thermal characteristics of strain-compensated GaInAsAlInAs quantum-cascade lasers", *Appl. Phys. Lett*. 91(16): 161111 (2007)

[12] M. Lindskog, J. M. Wolf, V. Trinite, v. Liverini, J. Faist, G. Maisons, M. Carras, R. Aidam, R. Ostendorf, and A. Wacker, "Comparative analysis of quantum cascade laser modeling based on density matrices and non-equilibrium Green's functions", *Appl. Phys. Lett.* **105,** 103106 (2014)

[13] S. Suri, Unpublished work.

[14] Y. Bai, N. Bandyopadhyay, S. Tsao, S. Slivken, and M. Razeghi, "Room temperature quantum cascade lasers with 27% wall plug efficiency", *Appl. Phys. Lett.* 98, 181102 (2011)

CHAPTER SEVEN

CONCLUSIONS

7.1 Conclusion

The over-arching goal of this work is to facilitate the creation of reliable and efficient QCLs. To achieve this, a multi-faceted approach has been employed as the reliability and efficiency of QCLs are tightly entwined. There a few novel results presented here to further this research:

1) Various methods of mitigating observed failure mechanisms, such as better thermal management techniques to avoid catastrophic mirror damage, progression from bonding devices with the soft solder indium, to a more reliable, CTE-matched, hard-solder AuSn, and the application of a dielectric overspray to passivate the HR-coating such that AuSn bonding does not lead to a damaged coating.
2) Correlation between thermal modeling for stationary and transient thermal responses performed with COMSOL Multiphysics and verified experimentally with CCD-based thermoreflectance measurements has been verified.
3) Preliminary results for an IFR-engineered device show the implementation of stepped interfaces at a few key positions can reduce the conduction band offset and push wavefunctions such that the wavefunctions' overlap is reduced. This has been shown for a specific design to reduce IFR leakage by 15% and increase the internal efficiency by 2%.
4) APT has been used to show the aluminum and gallium incorporation as a function of layer thickness. This has shown that thin layers, less than 2 nm in thickness, require a higher aluminum/gallium target to achieve an adequate amount of incorporation. This was

experimentally verified by increasing the molar ratio of Al in the gas flow during growth for thin (1.1 nm) layers which converged modeled emitting wavelength with experimental emitting wavelength.

5) APT has been used to extract IFR parameters from a full QCL structure emitting in the mid-IR for a few interfaces that play key roles in IFR scattering and leakage. It has been shown that barriers with a higher aluminum target, and in turn higher strain relative to the substrate, show larger rms roughness values when compared to barriers of shorter height, and in turn lower strain. The RMS roughness of the tallest barrier was found to be nearly 50% greater than that of the shorter barriers.
6) Application of the 3D characteristic of IFR scattering has been incorporated into a scattering model; this includes: the addition of the axial correlation length and interfacial mixing width, graded conduction band edge, quaternary AD scattering, graded lattice constant, and the use of the effective mass from each state instead of the effective mass at each point within the active region. The main result is significantly lower IFR leakage and a large reduction in LO transition efficiency. However, if the electronic temperatures are as high as is calculated NEGF software, the IFR leakage can increase by as much as an order of magnitude compared to what is presented.

With this work in place, a path forward to further improving device reliability and efficiency is presented through the use of a more accurate scattering model that incorporates the 3D aspect of layer gradings in QCLs.

7.2 Future Work

To further the work on reliable and efficient devices, reliability studies and APT analysis are key areas of interest. With the mitigation of many observed failures, further long-term, high-power reliability

studies would help identify internal failure mechanisms and extract activation energies. To eliminate potential oxidation of the front facet, devices can be run in a vacuum, a purged environment, or front facet optimization via sputtering can be investigate.

For a better understanding of the interfaces, interface roughness parameters from all of the layers within a full QCL stage can be analyzed and input into the scattering model for a more complete model. This can also be done for QCLs with varying growth conditions to optimize the growth conditions.

Further investigation of electronic temperatures are necessary to more accurately calculate IFR leakage currents.

Appendix I: Acronyms

AD – Alloy Disorder
APT – Atom Probe Tomography
AR – Anti-Reflective
BH – Buried Heterostructure
CB – Conduction Band
CCD – Charge-Coupled Device
COMD – Catastrophic Optical Mirror Damage
CTE – Coefficient of Thermal Expansion
CW – Continuous Wave
FIB – Focused Ion Beam
GSMBE – gas source molecular beam epitaxy
HAADF-STEM – High-Angle Annular Dark-Field Scanning Transmission Electron Microscopy
HHCF – Height-Height Correlation Function
HR – High-Reflectivity
HR-XRD – High Resolution X-Ray Diffraction
ICL – Interband Cascade Laser
IFR – Interface Roughness
LIV – Light-Current-Voltage
ll – lower laser
LO – Longitudinal Optical
LR – Low-Reflectivity
LUT – Look-Up Table
MBE – molecular beam epitaxy
MCT – Mercury-Cadmium-Telluride
Mid-IR – Mid-Infrared
MOCVD – Metal Organic Chemical Vapor Deposition
MTTF – Mean Time to Failure
NGEF – Non-Equilibrium Greene's Function
NRE – Nonresonant Extraction
QCL – Quantum Cascade Laser
QCW – Quasi-Continuous Wave
QW – Quantum Wells
RMS – Root Mean Square
ROI – Region of Interest
SEM – Scanning Electron Microscopy
SL – Superlattice
STA – Step-Tapered Active
STA-RE – Step-Tapered Active-Resonant Extraction
TA – Tapered-Active
TA-DW – Tapered Active-Deep Well
TEC – Thermo-Electric Cooler
TEM – Transmission Electron Microscopy
TMAl – Trimethylaluminum
TMGa – Trimethylgallium

TMIn – Trimethylindium
ul – upper laser
WPE – Wall-Plug Efficiency

Appendix II: APT Matlab Scripts

HHCF Script

```
clear all
close all

bin_width = 1; %decrease bin_width for accuracy (increases computation time)
            %bin_width = 1 --> ~1min, bin_width = 0.2 --> ~5min
percentage = 'Artificial Surface - Potato Chip#2';
%% 1) Import data
delimiter = ' ';
startRow =3;
formatSpec = '%f%f%f%f%f%f%f%[^\n\r]';

%put path to file in fileID 'C:\\ etc.... .txt'
%this will read .TXT files, .BCR must be converted to xyz data in Gwyddion
fileID = fopen('C:\Users\bbkni\Documents\Lab\2016-17\APT\Full
Structure\TallestBarrier_Isocentrations\Tall_Barrier_APT_Isoconcentration_BCR\Textfiles\TallestBarrier
_28per_towards_ActiveWell.txt','r');

dataArray = textscan(fileID, formatSpec, 'Delimiter', delimiter, 'EmptyValue' ,NaN,'HeaderLines'
,startRow, 'ReturnOnError', false);
fclose(fileID);
x = dataArray{:, 1}*1E9;%imports in meters, converting to nm
y = dataArray{:, 2}*1E9;
z = dataArray{:, 3}*1E9;

% meanvaluez = sum(z)/length(z);
% z = z-meanvaluez;

%% 2) Measure the distance and height difference between every combination of points
tic
q=1;
for i = 1:1:length(x)
   for j = i:1:length(x)
      distance(q) = sqrt(((x(i)-x(j))^2+(y(i)-y(j))^2)); %distance between 2 points (tau)
      H(q) = (abs(z(i)-z(j)))^2; %Difference in height between those 2 points squared
      q = q+1;
   end
end
Ht = vertcat(distance,H); % putting distance between two points and the squared difference in height to
an array
toc

%% 3) Binning and averaging all of the heights at their tau values
tic
```

```
for i = 1:1:length(Ht)
   if Ht(1,i) == 0 %tau (distance value)
   else
      Hholder = Ht(2,i); % abs(z(i) - z(j))^2
      avecount = 1;
      for j = i+1:1:length(Ht)
         if Ht(1,j) > Ht(1,i)-bin_width/2 && Ht(1,j) < Ht(1,i)+bin_width/2 %if the distance is between the
edges of the bin
            Hholder = Hholder + Ht(2,j); %if so, add (z(i)^2 - z(j)^2)^0.5
            avecount = avecount+1; %increase this count so we know how many are here to average
            Ht(1,j) = 0; % if it is averaged in, set to 0 so we can remove it otherwise this takes days to run
            Ht(2,j) = 0; % if it is averaged in, set to 0 so we can remove it
            %so if  Ht(1,j) = 0 line 40 will skip over it so it's not double
            %counted and significantly (100's of times) decreases the
            %runtime
         end
      end
      Ht(2,i) = Hholder/avecount; % average all the values within the bin width
   end
end
Ht( :, ~any(Ht,1) ) = [];  %switch rows to columns for plotting
toc


%% 4)Plotting and Fitting
figure(1)
fig1=figure(1);
plot(Ht(1,:),Ht(2,:),'o')
% ylim([0 0.4])
xlim([0 60])
hold on
xforfit = reshape(Ht(1,:),length(Ht),1);
yforfit = reshape(Ht(2,:),length(Ht),1);
delete_past_nm = 30;
rowsToDelete = xforfit > delete_past_nm; % choose how much of the data to include for the fit
xforfit(rowsToDelete) = [];
yforfit(rowsToDelete) = [];
xx = linspace(0,delete_past_nm)'; %how far the fit curve goes on the graph
startpoints = [0.2 3];
alpha = 1; % Hurst parameter, set to 1, when fit it comes out to ~0.97 for this interface
nonlinearfit = @(del,x)2*del(1).^2*(1-exp(-(((x)/del(2)).^(2*alpha))))
% nonlinearfit = @(del,x)2*0.18.^2*(1-exp(-(((x)/del(1)).^(2*alpha))))

w = 0;
for i = 1:1:length(xforfit) % This is to weight the fitting function,make the fit fit better to the beginning or
end of curve
```

```
   if i>50 && i <75
      w(i) = 1;
%    elseif i<37
%       w(i) = 1;
   else
      w(i) = 1;
   end
end
fitobject1 = fitnlm(xforfit,yforfit, nonlinearfit,startpoints,'weight',w)

figure(2)
fig2 = figure(2);
plot(xforfit,yforfit,'o') ;
hold on

caption = sprintf('H(\\tau) = 2\\Delta^{2}[1-exp(-(\\tau^{2}/\\Lambda^{2})] \n\\Delta = %1.3f nm,
\\Lambda = %1.2f nm \n            R^{2} = %0.3f',
fitobject1.Coefficients.Estimate(1),fitobject1.Coefficients.Estimate(2),fitobject1.Rsquared.Ordinary(1));
text(2,0.01,caption, 'FontSize', 14, 'FontWeight', 'bold');

line(xx,predict(fitobject1,xx),'color','b')
ylabel('H(\tau) (nm^{2})')
set(gca,'FontSize',20)
line(xx,predict(fitobject1,xx),'color','b')

del = 0.18;

lambda = 6.98;
xplot = linspace(0,30,1000);
yplot = 2*del^2*(1-exp(-(((xplot)./lambda).^2)));
% figure(3)
% plot(xplot,yplot,'--')
% ylim([0 .03])
xlabel('\tau (nm)')
ylabel('H(\tau) (nm^{2})')
title(percentage);
% xlim([0 15])
set(gca,'FontSize',20)
% xlim([0 10])
% filename = sprintf('%s',percentage);
% saveas(fig2,'35%-int684','tif')
% saveas(fig1,'Full-35%-int684','tif')
```

Interfacial Mixing Width Script

```
clc;
clear all;
close all;

csv = csvread('C:\Users\bbkni\Documents\Lab\2016-17\APT\1-2021\Jan-2021 NU
VISIT\5g_mod\Axial_correlation_Length\R5006_30595-v13_4thstagefrombottom.csv',1,0);

title = 'Stage 4, V13, Tall Barrier';
Al = csv(:,4);
In = csv(:,7);
Ga = csv(:,6);
As = csv(:,5);
Gapercent = Ga./(Ga+In+Al).*100;
Alpercent = Al./(Al+In+Ga).*100;
Inpercent = In./(Al+In+Ga).*100;
Aspercent = As;
Alpercent = Al;
X = csv(:,1);
Y = Alpercent;
plot(X,Y)

lowval = 18.1; %stage 2 First Barrier
upperlim = 20;

j = 1;
for i = 1:length(X)
   if X(i) <upperlim & X(i) >lowval
      holdx(j,1) = X(i)-lowval;
      holdy(j,1) = Alpercent(i);
      j = j+1;
   end
end
bin = 0.1;
lowerpad = 20;
if lowerpad == 0
   x(1,1) = 0;
   y(1,1) = 0;
end
for i = 1:(lowerpad)
   if i ==1
      x(i,1) = 0;
      y(i,1) = holdy(2,1);
   else
      x(i,1) = x(i-1,1) + bin;
```

```
      y(i,1) = holdy(2,1);
   end
end
i = length(lowerpad)+1+length(holdx);
rightpad = 20;
for h = 1:(rightpad)
   if h ==1
      x2(h,1)= holdx(j-1,1)+lowerpad*bin;
      y2(h,1) = holdy(j-1,1);
%       i = i+1;
   else
      x2(h,1) = x2(h-1,1)+bin;
      y2(h,1) = y2(h-1,1);
%       i = i+1;
   end
end
j = 1;
for i = 1:length(X)
   if X(i) <upperlim & X(i) >lowval
      holdx(j,1) = X(i)-lowval+lowerpad*bin;
      holdy(j,1) = Alpercent(i);
      j = j+1;
   end
end
x = cat(1,x,holdx,x2);
y = cat(1,y,holdy,y2);

figure(2)
plot(x,y)
hold on

hold on
xlabel('z (nm)')
ylabel('Aluminum Composition (%)')
modelfun=fittype('c0+d0*(erf(2*sqrt(log(2))*(x-e)/L))', 'independent', 'x', 'dependent', 'y' );

fitz = fit(x,y,modelfun,'StartPoint',[30, 27, 2, 1E-10])
% [mdl,GoF] = fit(x,y,modelfun,'StartPoint',[1,1, 1, 1]);

coeff = coeffvalues(fitz);
fitplot = coeff(2)+coeff(3)*erf(2*sqrt(log(2)).*(x-coeff(4))/coeff(1));
plot(x,fitplot)
caption = sprintf('%s \n Intefacial Width: %0.2fnm',title, abs(coeff(1)));
% text(156,172,caption, 'FontSize', 12, 'FontWeight', 'bold');
a = annotation('textbox',[0.15 0.7 0.4 0.2],'String',caption,'FitBoxToText','on','LineStyle','none');
a.FontSize = 12;
```

Axial Correlation Length Extraction Script

```
clear all
close all

%When the BCR file is exported the z component of the isoconcentration is
%lost.  This means the distance between interfaces is not contained within
%the output bcr files.


%% 1) Variables to Edit

folderlocation = 'C:\Users\bbkni\Documents\Lab\2016-17\APT\1-2021\Jan-2021 NU
VISIT\5g_mod\Axial_correlation_Length\3rd_stage\V15\Tall_Barrier\Correct_Barrier';

total_del = 0.253; %Stage 3 V15 Tall Barrier at Al = 37

titleName = 'Stage 4 V13 Tall Barrier';
bin_width = 0.05;

%% 2) Import data
delimiter = ' ';
startRow =3;
formatSpec = '%f%f%f%f%f%f%f%[^\n\r]';
txtpath = [folderlocation '\'];
txtfiles = dir([txtpath '*.txt']);
legendCell = {};

for k = 1:(size(txtfiles,1))
   fileID = fopen([txtpath txtfiles(k).name], 'r');
   dataArray = textscan(fileID, formatSpec, 'Delimiter', delimiter, 'EmptyValue' ,NaN,'HeaderLines'
,startRow, 'ReturnOnError', false);
   fclose(fileID);
   x(:,k) = dataArray{:, 1}*1E9;%imports in meters, converting to nm
   y(:,k) = dataArray{:, 2}*1E9;
   z(:,k) = dataArray{:, 3}*1E9;


end

meanvaluez = [5.9723887 5.991369 6.0129695 6.039991 6.0838833 6.10673 6.1286187 6.1548357
6.1780786]; % 3rd stage V15 Tall Barrier (centroids)


%% 3) Measure the distance between the average z-height of each interface and the respective axial
correlation length function
```

```
tic

a = 1;
q = 1;
g=1;
% for a = 1:1:(size(txtfiles,1))
for a = 4
   for b = 1:1:(size(txtfiles,1))
      for c = 1:1:length(x(:,1))
         for d = 1:1:length(x(:,1))
            if abs(x(c,a)-x(d,b)) < .1 && abs(y(c,a)-y(d,b)) < .1
               before_W_ij(a,b,q) = ((z(c,a)-z(d,b)))^2;
               before_x_ij(a,b,q) = x(c,a);
               before_y_ij(a,b,q) = y(c,a);

               otherperp(a,b,q) = z(c,a)*z(d,b);
%                 avecounter = avecounter+1;
               q = q+1;
            else
            end
         end
      end
      W_ij(a,b) = sqrt(mean(before_W_ij(a,b,:)));
      c_perp(a,b) = (1-W_ij(a,b).^2/(2*total_del^2));
      c_lin_perp(g) = c_perp(a,b);
      other_c_perp(g) = mean(otherperp(a,b,:))/total_del^2;
      del_z(a,b) = abs(meanvaluez(a)-meanvaluez(b));
      del_lin_z(g) = del_z(a,b);
      g = g+1;
      q = 1;


%       w = (1-(mean(sqrt((z(:,i)-z(:,j)).^2)))^2/(2*total_del^2));
%       if w > 100
%       else
%          w(i,j) = w;
%          del_z(i,j) = abs(meanvaluez(i)-meanvaluez(j));
%          w_hold(a) = w(i,j);
%          del_z_hold(a) = del_z(i,j);
%          a = a+1;
%       end

   end
end
```

```
scatter(del_lin_z,c_lin_perp)
hold on
scatter(del_lin_z,other_c_perp)
%%
a = 1;
% mean(W_ij)
% c_perp = 1-(mean(W_ij))^2/(2*0.169^2)
xforfit = del_lin_z;
yforfit = c_lin_perp;
rowsToDelete = xforfit > 10; % choose how much of the data to include for the fit
rowsToDelete = yforfit < 0; % choose how much of the data to include for the fit
xforfit(rowsToDelete) = [];
yforfit(rowsToDelete) = [];
xx = linspace(0,max(del_lin_z))'; %how far the fit curve goes on the graph
startpoints = [0.5];
alpha = 1; % Hurst parameter, set to 1, when fit it comes out to ~0.97 for this interface
nonlinearfit = @(del,x)(exp(-(x)/del(1)))

fitobject1 = fitnlm(xforfit,yforfit, nonlinearfit,startpoints)
figure(2)
fig2 = figure(2);
plot(xforfit,yforfit,'o') ;
hold on
line(xx,predict(fitobject1,xx),'color','b')

% ylim([0 1])


% figure(3)
% tri = delaunay(before_x_ij(1,3,:),before_y_ij(1,3,:),before_W_ij(1,3,:));
% surf(before_x_ij(1,3,:),before_y_ij(1,3,:),before_W_ij(1,3,:));
tic

caption = sprintf('C_{\\perp}(\\Deltaz) = exp(-\\Deltaz/\\Lambda_{\\perp}) \n      \\Lambda_{\\perp} =
%1.3f nm \n ', fitobject1.Coefficients.Estimate(1));
text(0.30,0.10,caption, 'FontSize', 14, 'FontWeight', 'bold');

line(xx,predict(fitobject1,xx),'color','b')
ylabel('C_{\perp}')
xlabel('\Deltaz (nm)')
set(gca,'FontSize',20)
line(xx,predict(fitobject1,xx),'color','b')
% plot(Ht(1,:),Ht(2,:),'o','color','g')
% plot(xval,yval,'o','color','r')
title(titleName)
ylim([0 1])
```

```
xlim([0 0.5])
hold on
```